

541,993

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
29. Juli 2004 (29.07.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/063358 A1

(51) Internationale Patentklassifikation⁷: C12N 1/15,
15/80, C12P 23/00

(21) Internationales Aktenzeichen: PCT/EP2004/000100

(22) Internationales Anmeldedatum:
9. Januar 2004 (09.01.2004)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
103 00 649.4 9. Januar 2003 (09.01.2003) DE
103 41 272.7 8. September 2003 (08.09.2003) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme
von US): BASF AKTIENGESELLSCHAFT [DE/DE];
67056 Ludwigshafen (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): MATUSCHEK,
Markus [DE/DE]; Karolinenstr. 5, 69469 Weinheim
(DE). HEINEKAMP, Thorsten [DE/DE]; Alte Ziegelei
1b, 38419 Hannover (DE). SCHMIDT, Andre [DE/DE];
Magdeburger Str. 11, 31832 Springe (DE). BRAKHAGE,
Axel [DE/DE]; Schneiderberg 58, 38167 Hannover (DE).

(74) Anwalt: FITZNER, Uwe; Lintorfer Strasse 10, 40878
Ratingen (DE).

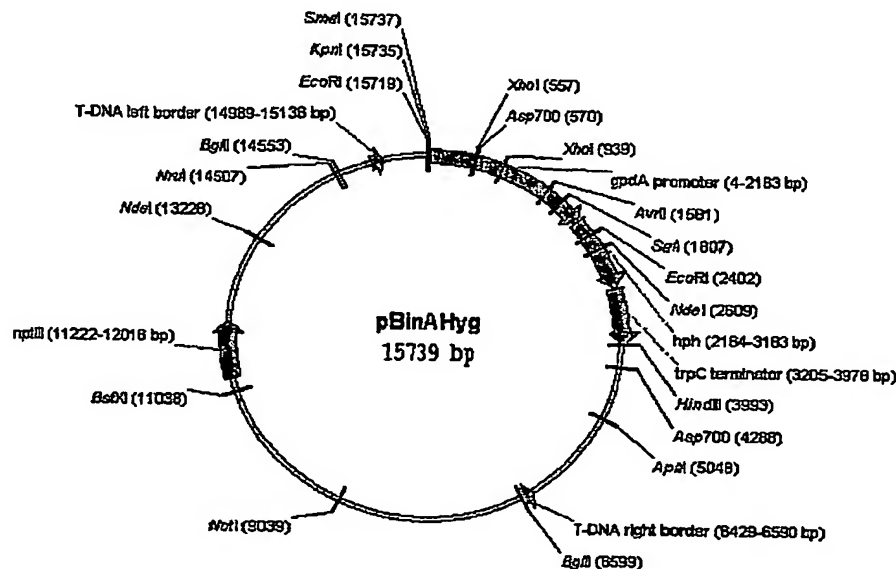
(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH,
CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR THE GENETIC MODIFICATION OF ORGANISMS OF THE GENUS BLAKESLEA, CORRE-
SPONDING ORGANISMS, AND THE USE OF THE SAME

(54) Bezeichnung: VERFAHREN ZUR GENTECHNISCHEN VERÄNDERUNG VON ORGANISMEN DER GATTUNG BLA-
KESLEA, ENTSPRECHENDE ORGANISMEN UND DEREN VERWENDUNG

VEKTOR pBinAHyg



(57) Abstract: The invention relates to a method for producing a genetically modified organism of the genus Blakeslea, said method comprising the following steps: (i) at least one of the cells is transformed, (ii) the cells obtained in step (i) are optionally rendered homokaryotic, so that cells are created in which the nuclei are all homogeneously modified in at least one genetic characteristic and convert said genetic modification into an expression, and (iii) the genetically modified cell or cells are selected and cultivated.

[Fortsetzung auf der nächsten Seite]

WO 2004/063358 A1



PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TI, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

— vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen

(84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

Veröffentlicht:

— mit internationalem Recherchenbericht

(57) **Zusammenfassung:** Verfahren zur gentechnischen Veränderung von Organismen der Gattung *Blakeslea*, entsprechende Organismen und deren Verwendung. Zusammenfassung Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* umfassend (i) Transformation mindestens einer der Zellen, (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder mehreren genetischen Merkmalen alle gleichartig verändert sind und diese gentechnische Veränderung zur Ausprägung bringen, und (iii) Selektion und Anzucht der gentechnisch veränderten Zelle oder Zellen.

**Verfahren zur gentechnischen Veränderung von Organismen der
Gattung Blakeslea, entsprechende Organismen und deren
Verwendung**

Die Erfindung betrifft ein Verfahren zur gentechnischen Veränderung von Organismen der Gattung Blakeslea, entsprechende Organismen und deren Verwendung.

Pilze der Gattung Blakeslea sind als Produktionsorganismen bekannt. So wird z. B. Blakeslea trispora als Produktionsorganismus für β -Carotin (Ciegler, 1965, Adv Appl Microbiol. 7:1) und Lycopin verwendet (EP 1201762, EP 1184464, WO 03/038064). Daneben kommt Blakeslea zur Produktion anderer lipophiler Substanzen in Frage wie z.B. andere Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder zur Produktion hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitamine, Provitamine und Cofaktoren.

Die hohen Produktivitäten für β -Carotin und Lycopin machen Blakeslea, insbesondere Blakeslea trispora attraktiv für die wirtschaftliche fermentative Herstellung von Carotinoiden und deren Vorstufen.

Allerdings ist es auch von Interesse, die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle zu ermöglichen, die von Blakeslea bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Carotinoide werden Futtermitteln, Nahrungsmitteln, Nahrungsergänzungsmitteln, Kosmetika und Arzneimitteln zugesetzt. Die Carotinoide dienen vor allem als Pigmente zur Färbung. Daneben werden

die antioxidative Wirkung der Carotinoide und andere Eigenschaften dieser Substanzen genutzt. Man unterteilt die Carotinoide in die reinen Kohlenwasserstoffe, die Carotine und die sauerstoffhaltigen Kohlenwasserstoffe, die Xanthophylle. Xanthophylle wie Canthaxanthin und Astaxanthin werden beispielsweise zur Pigmentierung von Hühnereiern und Fischen eingesetzt (Britton et al. 1998, Carotinoids, Vol 3, Biosynthesis and Metabolism). Die Carotine β -Carotin und Lycopin werden vor allem in der Humanernährung eingesetzt. β -Carotin wird beispielsweise als Getränkfarbstoff verwendet. Lycopin hat eine krankheitsvorbeugende Wirkung (Argwal und Rao, 2000, CMAJ 163:739-744; Rao und Argwal 1999, Nutrition Research 19:305-323). Die farblose Carotinoidvorstufe Phytoen kommt vor allem für Anwendungen als Antioxidans in Frage.

Der überwiegende Teil der Carotinoide und deren Vorstufen, die als Zusatzstoffe für die oben genannten Anwendungen eingesetzt werden, wird durch chemische Synthese hergestellt. Die chemische Synthese ist mehrstufig, technisch sehr aufwendig und verursacht hohe Herstellkosten. Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten.

Ein Verfahren zur gentechnischen Veränderung von *Blakeslea trispora* ist erforderlich insbesondere, wenn *Blakeslea* zur Herstellung von Xanthophyllen genutzt werden soll, weil diese Verbindungen natürlicherweise von *Blakeslea* nicht synthetisiert werden.

Von *Blakeslea trispora* sind bisher verschiedene DNA-Sequenzen bekannt insbesondere die DNA-Sequenz, die für die Gene der Carotinoidbiosynthese von Geranylgeranylpyrophosphat bis β -Carotin codiert (WO 03/027293).

5

Allerdings sind bisher keine Methoden zur gentechnischen Veränderung von *Blakeslea*, insbesondere *Blakeslea trispora* bekannt.

Als Methode zur Herstellung von gentechnischen veränderten Pilzen wurde in einigen Fällen die *Agrobacterium*-vermittelte Transformation erfolgreich eingesetzt. So sind z. B. folgende Organismen durch Agrobakterien transformiert worden: *Saccharomyces cerevisiae* (Bundock et al., 1995, EMBO Journal, 14:3206–3214), *Aspergillus awamori*, *Aspergillus nidulans*, *Aspergillus niger*, *Colletotrichum gloeosporioides*, *Fusarium solani pisi*, *Neurospora crassa*, *Trichoderma reesei*, *Pleurotus ostreatus*, *Fusarium graminearum* (van der Toorren et al., 1997, EP 870835), *Agraricus bisporus*, *Fusarium venenatum* (de Groot et al., 1998, Nature Biotechnol. 16:839–842), *Mycosphaerella graminicola* (Zwiers et al. 2001, Curr. Genet. 39:388–393), *Glarea lozoyensis* (Zhang et al., 2003, Mol. Gen. Genomics 268:645–655), *Mucor miehei* (Monfort et al. 2003, FEMS Microbiology Lett. 244:101 – 106).

Von Interesse ist besonders eine homologe Rekombination, bei der zwischen der einzuführenden DNA und der Zell-DNA möglichst viele Sequenzhomologien bestehen, so dass eine ortsspezifische Einführung bzw. Ausschaltung von genetischer Information im Genom des Empfängerorganismus möglich ist. Andernfalls wird die Spender-DNA durch illegitime bzw. nicht-homologe Rekombination ins Genom des Empfängerorganismus integriert, was nicht ortsspezifisch erfolgt.

Eine durch *Agrobacterium* vermittelte Transformation und anschließende homologe Rekombination der transferierten DNA wurde bisher bei folgenden Organismen nachgewiesen: *Aspergillus awamori* (Gouka et al. 1999, *Nature Biotech* 17:598-601), *Glarea lozoyensis* (Zhang et al., 2003, Mol. Gen. Genomics 268:645-655), *Mycosphaerella graminicola* (Zwiers et al. 2001, *Curr. Genet.* 39:388-393).

Als weitere Methode zur Transformation von Pilzen ist die Elektroporation bekannt. Die integrative Transformation von Hefe durch Elektroporation wurde von Hill, *Nucl. Acids. Res.* 17:8011 gezeigt. Für filamentöse Pilze wurde die Transformation durch Chakabarty und Kapoor beschrieben (1990, *Nucl. Acids. Res.* 18:6737).

Eine „biolistische“ Methode, d.h. die Übertragung von DNA durch Beschuss von Zellen mit DNA-beladenen Partikeln wurde beispielsweise für *Trichoderma harzianum* und *Gliocladium virens* beschrieben (Lorito et al. 1993, *Curr. Genet.* 24:349–356).

Diese Methoden konnten bisher jedoch nicht erfolgreich zur gezielten genetischen Veränderung von *Blakeslea* und insbesondere *Blakeslea trispora* eingesetzt werden.

Eine besondere Schwierigkeit bei der Herstellung von gezielt genetisch veränderten *Blakeslea* und *Blakeslea trispora* ist die Tatsache, dass deren Zellen in allen Stadien des sexuellen und des vegetativen Zellzyklus mehrkernig sind. In Sporen von *Blakeslea trispora* Stamm NRRL2456 und NRRL2457 wurden z. B. im Durchschnitt 4,5 Kerne pro Spore nachgewiesen (Metha und Cerdá-Olmedo, 1995, *Appl. Microbiol. Biotechnol.* 42:836–838). Dies hat zur Folge, dass die gentechnische Veränderung in aller Regel nur in einem oder wenigen Kernen vorliegt, die Zellen also heterokaryotisch sind.

Sollen die genetisch veränderten Blakeslea-Arten, insbesondere Blakeslea trispora zur Produktion eingesetzt werden, so ist es insbesondere bei einer Gendeletion wichtig, dass in den
5 Produktionsstämmen die gentechnische Veränderung in allen Kernen vorliegt, so dass eine stabile und hohe Syntheseleistung ohne Nebenprodukte möglich wird. Die Stämme müssen folglich in Bezug auf die gentechnische Veränderung homokaryotisch sein.

10 Lediglich für Phycomyces blakesleeanus ist ein Verfahren beschrieben worden, um homokaryotische Zellen zu erzeugen (Roncero et al., 1984, Mutat. Res. 125:195). Durch Zugabe des mutagenen Agens MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) werden nach dem dort beschriebenen Verfahren Kerne in den Zellen eliminiert, so dass statistisch eine gewisse
15 Anzahl von Zellen mit nur noch einem funktionellem Kern vorliegt. Die Zellen werden dann einer Selektion unterzogen, in der nur einkernige Zellen mit einem rezessiven Selektionsmarker zu einem Mycel auswachsen können. Die Nachkommen dieser selektierten Zellen sind mehrkernig und homokaryotisch. Ein rezessiver Selektionsmarker für
20 Phycomyces blakesleeanus ist z. B. dar. Dar^+ -Stämme nehmen das toxische Riboflavin-Analog 5-Carbon-5-deazariboflavin auf; Dar^- -Stämme dagegen nicht (Delbrück et al. 1979, Genetics 92:27). Die Selektion von rezessiven Mutanten erfolgt durch Zugabe von 5-Carbon-5-deazariboflavin (DARF).

25

Allerdings ist dieses Verfahren nicht für Blakeslea, insbesondere Blakeslea trispora bekannt und insbesondere nicht mit im Zusammenhang mit einer Transformation beschrieben worden.

30 Aufgabe der vorliegenden Erfindung ist es, ein Verfahren bereitzustellen, mit dem eine gentechnische Veränderung von Blakeslea-Stämmen,

insbesondere *Blakeslea trispora* möglich ist. Darüber hinaus ist es Aufgabe der Erfindung ein Verfahren bereitzustellen, das die Herstellung homokaryotischer genetisch veränderter Stämme erlaubt. Ferner ist es eine Aufgabe der Erfindung entsprechend gentechnisch veränderte Zellen
5 bereitzustellen.

Diese Aufgabe wird durch ein Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* gelöst, umfassend

- 10 (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und
- 15 (iii) Selektion der gentechnisch veränderten Zelle oder Zellen.

Mit der erfindungsgemäßen Methode ist es möglich, mehrkernige Zellen der Pilze *Blakeslea* gezielt und stabil genetisch zu verändern, um so Mycel aus Zellen mit einheitlichen Kernen zu gewinnen. Vorzugsweise handelt
20 es sich Zellen von Pilzen der Art *Blakeslea trispora*.

Unter Transformation wird die Übertragung einer genetischen Information in den Organismus, insbesondere Pilz verstanden. Darunter sollen alle dem Fachmann bekannten Möglichkeiten zur Einschleusung der
25 Information, insbesondere DNA fallen, z. B. Beschuss mit DNA-beladenen Partikeln, Transformation mittels Protoplasten, Mikroinjektion von DNA, Elektroporation, Konjugation oder Transformation kompetenter Zellen, Chemikalien oder Agrobakterien vermittelte Transformation. Als genetische Information werden ein Genabschnitt, ein Gen oder mehrere

Gene verstanden. Die genetische Information kann z. B. mit Hilfe eines Vectors oder als freie Nukleinsäure (z. B. DNA, RNA) und auf sonstige Weise in die Zellen eingebracht und entweder durch Rekombination ins Wirtsgenom eingebaut oder in freier Form in der Zelle vorliegen.

5 Besonders bevorzugt ist hierbei die homologe Rekombination.

Bevorzugte Transformationsmethode ist die *Agrobacterium tumefaciens*-vermittelte Transformation. Hierzu wird zunächst die zu transferierende Spender-DNA in einen Vektor eingefügt, der (i) flankierend zu der zu
10 transferierenden DNA die T-DNA-Enden trägt, der (ii) einen Selektionsmarker enthält und der (iii) ggf. Promotoren und Terminatoren für die Genexpression der Spender-DNA aufweist. Dieser Vektor wird in einen *Agrobacterium tumefaciens*-Stamm übertragen, der ein Ti-Plasmid mit den vir-Genen enthält. vir-Gene sind für den DNA-Transfer in
15 *Blakeslea* verantwortlich. Mit diesem Zwei-Vektor-System wird die DNA von *Agrobacterium* in *Blakeslea* übertragen. Hierzu werden die Agrobakterien zunächst in Gegenwart von Acetosyringone inkubiert. Acetosyringone induziert die vir-Gene. Anschließend werden Sporen von *Blakeslea trispora* zusammen mit den induzierten Zellen von
20 *Agrobacterium tumefaciens* auf Acetosyringone-haltigem Medium inkubiert und dann auf Medium übertragen, das eine Selektion der Transformanten, d.h. der gentechnisch veränderten Stämme von *Blakeslea* ermöglicht.

Der Begriff Vector wird in der vorliegenden Anmeldung als eine
25 Bezeichnung für ein DNA-Molekül verwendet, das zum Einschleusen und ggf. zur Vermehrung von Fremd-DNA in eine Zelle dient (siehe auch "Vector" in Römpp Lexikon Chemie – CDRom Version 2.0, Stuttgart/New York: Georg Thieme Verlag 1999). In der vorliegenden Anmeldung sollen unter dem Begriff "Vector" Plasmide, Cosmide usw. verstanden werden,
30 die diesem Zweck dienen.

Unter Expression wird in der vorliegenden Anmeldung die Übertragung einer genetischen Information ausgehend von DNA oder RNA in ein Gen-Produkt (hier vorzugsweise Carotinoide) verstanden und soll auch den Begriff der Überexpression beinhalten, womit eine verstärkte Expression gemeint ist, so dass ein bereits in der nicht transformierten Zelle (Wildtyp) hergestelltes Produkt verstärkt produziert wird oder einen großen Teil des gesamten Gehaltes der Zelle ausmacht.

Unter gentechnischer Veränderung soll die Einschleusung genetischer Information in einen Empfängerorganismus, so dass diese stabil exprimiert und bei der Zellteilung weitergegeben wird, verstanden werden. Danach wird gegebenenfalls die Homokaryontisierung durchgeführt, d.h. die Herstellung von Zellen, die nur einheitliche Kerne enthalten, d. h. Kerne mit gleichem genetischem Informationsgehalt.

Diese Homokaryontisierung ist insbesondere notwendig, wenn die durch Transformation eingeführte genetische Information rezessiv vorliegt, d. h. nicht zur Ausprägung gelangt. Führt die Transformation aber zu einem dominanten Vorliegen der genetischen Information, d. h. wird sie ausgeprägt, so ist eine Homokaryontisierung nicht unbedingt nötig.

Vorzugsweise wird zur Homokaryontisierung eine Selektion der einkernigen Sporen durchgeführt. Von Natur aus ist ein geringer Anteil der Sporen von *Blakeslea trispora* einkernig, so dass sich diese ggf. nach spezifischer Markierung z. B. Färbung der Zellkerne aussortieren lassen. Dies wird bevorzugterweise mittels FACS (Fluorescence Activated Cell Sorting) anhand der geringeren Fluoreszenz der einkernigen Zellen durchgeführt.

Alternativ kann zur Homokaryontisierung zunächst eine Kernreduktion durchgeführt werden. Hierzu kann ein mutagenes Agens eingesetzt werden, wobei es sich insbesondere um N-Methyl-N'-nitro-nitrosoguanidin

(MNNG) handelt. Auch die Verwendung von energiereichen Strahlen, wie UV- oder Röntgen-Strahlen zur Kernreduktion ist möglich. Anschließend kann zur Selektion auf das FACS Verfahren oder rezessive Selektionsmarker zurückgegriffen werden.

5

Unter Selektion wird die Auswahl von Zellen verstanden, deren Kerne dieselbe genetische Information beinhalten, d. h. Zellen die die gleichen Eigenschaften aufweisen, wie Resistenzen oder die Herstellung bzw. vermehrte Herstellung eines Produktes. In der Selektion werden neben
10 der FACS Methode bevorzugt 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) oder 5'-Fluororotat (FOA) und Uracil eingesetzt.

Der in der Transformation (i) eingesetzte Vector kann derart gestaltet sein, dass die im Vector enthaltene genetische Information in das Genom
15 mindestens einer Zelle integriert wird. Dabei kann genetische Information in der Zelle ausgeschaltet werden.

Der in der Transformation (i) eingesetzte Vector kann aber auch derart ausgestaltet sein, dass die im Vector enthaltene genetische Information in
20 der Zelle exprimiert wird, d. h. genetische Information eingefügt wird, die im korrespondierenden Wildtyp nicht vorhanden ist oder die durch die Transformation verstärkt bzw. überexprimiert wird.

Der Vector kann beliebige genetische Informationen zur genetischen
25 Veränderungen von Organismen der Gattung *Blakeslea* enthalten.

Unter „genetischer Information“ werden vorzugsweise Nukleinsäuren verstanden, deren Einbringung in den Organismus der Gattung *Blakeslea* zu einer genetischen Veränderung in Organismen der Gattung *Blakeslea*,
30 also beispielsweise zu einer Verursachung, Erhöhung oder Reduzierung von Enzymaktivitäten im Vergleich zum Ausgangsorganismus führen.

Der Vector kann beispielsweise genetische Information zur Herstellung lipophiler Substanzen enthalten wie z.B. Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder genetische Information zur Herstellung hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitaminen, Provitamine und Cofaktoren.

Bevorzugterweise enthält der eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder Xanthophyllen oder deren Vorstufen.

Bevorzugterweise enthält der Vektor genetische Information, die eine Lokalisierung der Carotinoidbiosynthese-Enzyme in dem Zellkompartiment bewirkt, in dem die Carotinoidbiosynthese stattfindet.

Besonders bevorzugt sind genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Adonixanthin, Adonirubin, Canthaxanthin, 3- und 3'-Hydroxyechinenon, Lycopin, Lutein, β -Carotin, Phytoen oder Phytoflu. Ganz besonders bevorzugt sind genetische Informationen zur Herstellung von Phytoen, Bixin, Lycopin, Zeaxanthin, Canthaxanthin und Astaxanthin.

Entsprechend werden in einer bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die über eine erhöhte Syntheserate für Zwischenprodukte der Carotinoidbiosynthese verfügen und folglich eine erhöhte Produktivität für Endprodukte der Carotinoidbiosynthese aufweisen. Zur Erhöhung der Syntheserate für Zwischenprodukte der Carotinoidbiosynthese werden insbesondere die Aktivitäten der Enzyme 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase,

Isopentenylpyrophosphat-Isomerase und Geranylpyrophosphatsynthase gesteigert.

Entsprechend werden in einer besonders bevorzugten Variante der
5 Erfindung Organismen hergestellt und kultiviert, die gegenüber dem Wildtyp eine erhöhte HMG-CoA-Reduktase-Aktivität aufweisen.

Unter HMG-CoA-Reduktase-Aktivität wird die Enzymaktivität einer HMG-CoA-Reduktase (3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase)
10 verstanden.

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

15 Dementsprechend wird unter HMG-CoA-Reduktase-Aktivität die in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. gebildete Menge Mevalonat verstanden.

20 Bei einer erhöhten HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase die umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. die gebildete Menge Mevalonat erhöht.

25 Vorzugsweise beträgt diese Erhöhung der HMG-CoA-Reduktase-Aktivität mindestens 5%, weiter bevorzugt mindestens 20%, weiter bevorzugt mindestens 50%, weiter bevorzugt mindestens 100%, bevorzugter mindestens 300%, noch bevorzugter mindestens 500%, insbesondere mindestens 600% der HMG-CoA-Reduktase-Aktivität des Wildtyps.

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp durch eine Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase.

- 5 In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens erfolgt die Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase indem man ein Nukleinsäurekonstrukt, enthaltend eine Nukleinsäure codierend eine HMG-CoA-Reduktase in den Organismus einbringt, deren Expression in
10 dem Organismus, verglichen mit dem Wildtyp, einer reduzierten Regulation unterliegt.

Unter einer reduzierten Regulation verglichen mit dem Wildtyp, wird eine im Vergleich zum vorstehend definierten Wildtyp verringerte, vorzugsweise
15 keine Regulation auf Expressions- oder Proteinebene verstanden.

Die reduzierte Regulation kann vorzugsweise durch einen im Nukleinsäurekonstrukt mit der kodierenden Sequenz funktionell verknüpften Promotor erreicht werden, der in dem Organismus verglichen
20 mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

Beispielsweise unterliegen die Promotoren ptef1 aus *Blakeslea trispora* und pgpdA aus *Aspergillus nidulans* nur einer reduzierten Regulation und sind daher insbesondere als Promotoren bevorzugt.

25

Diese Promotoren zeigen eine annähernd konstitutive Expression in *Blakeslea trispora*, so dass die transkriptionelle Regulation nicht mehr über die Intermediate der Carotinoidbiosynthese abläuft.

- 30 Die reduzierte Regulation kann in einer weiteren bevorzugten Ausführungsform dadurch erreicht werden, dass man als Nukleinsäure

codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt.

5 Besonders bevorzugt ist die Verwendung einer Nukleinsäure, die nur den katalytischen Bereich der HMG-CoA-Reduktase kodiert (trunkierte (t-)HMG-CoA-Reduktase). Die für die Regulation verantwortliche Membran-Domäne fehlt. Die verwendete Nukleinsäure unterliegt somit einer reduzierten Regulation und führt zu einer Erhöhung der Genexpression
10 der HMG-CoA-Reduktase. In einer besonders bevorzugten Ausführungsform bringt man Nukleinsäuren in *Blakeslea trispora* ein, welche die Sequenz SEQ ID. NO. 75 enthalten.

15 Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Homologievergleiche der Sequenzen aus Datenbanken mit der SEQ ID.
20 NO. 75 leicht auffinden.

Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich weiterhin beispielsweise ausgehend von
25 der Sequenz SEQ ID. NO. 75 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

In einer besonders bevorzugten Ausführungsform wird die reduzierte
30 Regulation dadurch erreicht, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in

dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt und einen Promotor verwendet, der in dem Organismus, verglichen mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

5

Entsprechend wird in einer bevorzugten Variante der Erfindung durch die Transformation die Genexpression der Phytoendesaturase ausgeschaltet, so dass das von den Organismen produzierte Phytoen gewonnen werden kann. Der in der Transformation (i) eingesetzte Vector umfasst daher in
10 einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Phytoendesaturase, insbesondere carB aus *Blakeslea trispora* mit der SEQ ID NO: 69.

Entsprechend wird in einer bevorzugten Variante der Erfindung durch
15 Transformation die Genexpression der Lycopincyclase ausgeschaltet, so dass das von den Organismen produzierte Lycopin gewonnen werden kann. Der in der Transformation eingesetzte Vektor umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Lycopincyclase, insbesondere
20 carR aus *Blakeslea trispora* s. (WO 03/027293).

In einer weiteren bevorzugten Ausführungsform werden die Organismen der Gattung *Blakeslea* beispielsweise in die Lage versetzt Xanthophylle, wie beispielsweise Zeaxanthin oder Astaxanthin herzustellen, indem die
25 genetisch veränderten Organismen der Gattung *Blakeslea* im Vergleich zum Wildtyp eine Hydroxylase-Aktivität und/oder eine Ketolase-Aktivität besitzen.

Der in der Transformation (i) eingesetzte Vector enthält also in einer
30 weiteren, bevorzugten Variante der Erfindung genetische Informationen, die nach Expression eine Ketolase- und/oder Hydroxylase-Aktivität

entfalten, so dass die Organismen Zeaxanthin oder Astaxanthin produzieren.

Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase
5 verstanden.

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -Ionon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

10

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Canthaxanthin umzuwandeln.

15 Dementsprechend wird unter Ketolase-Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge β -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

Unter dem Begriff "Wildtyp" wird erfindungsgemäß der entsprechende
20 nicht genetisch veränderte Ausgangsorganismus der Gattung *Blakesleaa* verstanden.

Je nach Zusammenhang kann unter dem Begriff "Organismus" der Ausgangsorganismus (Wildtyp) der Gattung *Blakesleaa* oder ein
25 erfindungsgemäßer, genetisch veränderter Organismus der Gattung *Blakesleaa* oder beides verstanden werden.

Vorzugsweise wird unter "Wildtyp" für die Verursachung der Ketolase-Aktivität und für die Verursachung der Hydroxylase-Aktivität jeweils ein
30 Referenzorganismus verstanden.

Dieser Referenzorganismus der Gattung *Blakeslea* ist *Blakeslea trispora* ATCC 14271 oder ATCC 14272, die sich lediglich im Paarungstyp unterscheiden.

- 5 Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakesleaa* und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Bestimmung der Ketolase-Aktivität in Organismen der Gattung
10 *Blakeslea* erfolgt in Anlehnung an die Methode von Fraser et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die Ketolase-Aktivität in Extrakten wird mit den Substraten beta-Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumcholat) bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels
15 HPLC ermittelt.

Der erfindungsgemäße genetisch veränderte Organismus der Gattung *Blakesleaa* weist in dieser, bevorzugten Ausführungsform im Vergleich zum genetisch nicht veränderten Wildtyp eine Ketolase-Aktivität auf und ist
20 somit vorzugsweise in der Lage, transgen eine Ketolase zu exprimieren.

In einer weiter bevorzugten Ausführungsform erfolgt die Verursachung der Ketolase-Aktivität in den Organismen der Gattung *Blakesleaa* durch Genexpression einer Nukleinsäure kodierend eine Ketolase.
25

In dieser bevorzugten Ausführungsform erfolgt die Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren in die Ausgangsorganismus der Gattung *Blakesleaa*.
30

Dazu kann prinzipiell jedes Ketolase-Gen, also jede Nukleinsäuren die eine Ketolase codiert verwendet werden.

Alle in der Beschreibung erwähnten Nukleinsäuren können beispielsweise
5 eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der Wirtsorganismus der Gattung Blakesleaa nicht in der Lage ist oder nicht in die Lage versetzt werden
10 kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

Beispiele für Nukleinsäuren, kodierend eine Ketolase und die
15 entsprechenden Ketolasen, die im erfindungsgemäßen Verfahren verwendet werden können sind beispielsweise Sequenzen aus:

Haematoccus pluvialis, insbesondere aus Haematoccus pluvialis Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 11, Protein
20 SEQ ID NO: 12),

Haematoccus pluvialis, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 13, Protein SEQ ID NO: 14),

25 Agrobacterium aurantiacum (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),

Aliccaligenes spec. (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),
30

Paracoccus marcusii (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20),

5 Synechocystis sp. Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22),

Bradyrhizobium sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24),

10 Nostoc sp. Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26),

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No. NZ_AABC01000195, Basenpaar 55,604 bis 55,392 (SEQ ID NO: 27);
15 Protein: Acc.-No. ZP_00111258 (SEQ ID NO: 28) (als putatives Protein annotiert) oder

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No. NZ_AABC01000196, Basenpaar 140,571 bis 139,810 (SEQ ID NO: 29),
20 Protein: (SEQ ID NO: 30) (nicht annotiert).

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im erfindungsgemäßen Verfahren verwendet werden können, lassen sich beispielsweise aus verschiedenen Organismen, deren genomische
25 Sequenz bekannt ist, durch Identitätsvergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen und insbesondere mit den Sequenzen SEQ ID NO: 12 und/oder 26 und/oder 30 leicht auffinden.

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 12 und/oder 26 und/oder 30 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenten (hohe Stringenz) Bedingungen erfolgen.

Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben.

Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit 2X SSC bei 50°C) und solchen mit hoher Stringenz (mit 0,2X SSC bei 50°C, bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).

Darüberhinaus kann die Temperatur während des Waschschrilles von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.

Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS

eingesetzt werden. In Gegenwart von 50 % Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt.

Einige beispielhafte Bedingungen für Hybridisierung und Waschschrift sind
5 infolge gegeben:

- (1) Hybridisierungsbedingungen mit zum Beispiel
 - (i) 4X SSC bei 65°C, oder
 - (ii) 6X SSC bei 45°C, oder
 - 10 (iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder
 - (iv) 6X SSC, 0,5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA bei 68°C, oder
 - (v) 6XSSC, 0,5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA, 50 % Formamid bei 42°C, oder
 - 15 (vi) 50 % Formamid, 4X SSC bei 42°C, oder
 - (vii) 50 % (vol/vol) Formamid, 0,1 % Rinderserumalbumin, 0,1 % Ficoll, 0,1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6,5, 750 mM NaCl, 75 mM Natriumcitrat bei 42°C, oder
 - (viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder
 - 20 (ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (moderate Bedingungen).

- (2) Waschschriffe für jeweils 10 Minuten mit zum Beispiel
 - (i) 0,015 M NaCl/0,0015 M Natriumcitrat/0,1 % SDS bei 50°C, oder
 - 25 (ii) 0,1X SSC bei 65°C, oder
 - (iii) 0,1X SSC, 0,5 % SDS bei 68°C, oder
 - (iv) 0,1X SSC, 0,5 % SDS, 50 % Formamid bei 42°C, oder
 - (v) 0,2X SSC, 0,1 % SDS bei 42°C, oder
 - (vi) 2X SSC bei 65°C (moderate Bedingungen).

In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch veränderten Organismen der Gattung *Blakeslea* bringt man Nukleinsäuren ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 12 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 %, insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz z SEQ ID NO: 12 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 12 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 26 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 % insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz SEQ ID NO: 26 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 26 durch
5 künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen
10 Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 30 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, bevorzugter mindestens 40 %, 15 bevorzugter mindestens 50 %, bevorzugter mindestens 60 %, bevorzugter mindestens 70 %, bevorzugter mindestens 80 %, besonders bevorzugt mindestens 90 %, insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz SEQ ID NO: 30 und die enzymatische Eigenschaft einer Ketolase aufweist.

20

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 30 durch
25 künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren
30 zu verstehen. Bevorzugt werden sog. konservative Austausche durchgeführt, bei denen die ersetzte Aminosäure eine ähnliche

Eigenschaft hat wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

- 5 Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung. Bevorzugte Positionen für Deletionen sind die Termini des Polypeptids und die Verknüpfungen zwischen den einzelnen Proteindomänen.

10 Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren ersetzt wird.

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinlänge verstanden, insbesondere die
15 Identität die durch Vergleich mit Hilfe der Lasergene Software der Firma DNASTAR, inc. Madison, Wisconsin (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

20

Multiple alignment parameter:

Gap penalty 10

Gap length penalty 10

Pairwise alignment parameter:

25 K-tuple 1

Gap penalty 3

Window 5

Diagonals saved 5

- 30 Unter einem Protein, das eine Identität von mindestens 20 % auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 oder 26 oder 30

aufweist, wird dementsprechend ein Protein verstanden, das bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 12 oder 26 oder 30, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 20 %, bevorzugt 80, %, 85%,
5 besonders 90%, insbesondere 95% aufweist.

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

10

Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Blakesleaaaspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene von Organismen der Gattung Blakesleaa leicht ermitteln.

15

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 11 in die Organismus der Gattung ein.

20 In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 25 in die Organismus der Gattung ein.

In einer weiteren, besonders bevorzugten Ausführungsform bringt man
25 eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 29 in die Organismus der Gattung ein.

Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen
30 wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die

chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoramiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des
5 Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

10 Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für eine Ketolase, insbesondere der Ketolase *Nostoc punctiforme* aus mit der SEQ ID NO: 72.

15 Unter Hydroxylase-Aktivität die Enzymaktivität einer Hydroxylase verstanden.

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -
20 Ionon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Zeaxanthin oder Cantaxanthin in Astaxanthin umzuwandeln.

25

Dementsprechend wird unter Hydroxyase-Aktivität die in einer bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge β -Carotin oder Cantaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

30

Bei einer erhöhten Hydroxylase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein Hydroxylase die umgesetzte Menge β -Carotin oder Canthaxantin bzw. die gebildete Menge Zeaxanthin oder Astaxanthin erhöht.

5

Vorzugsweise beträgt diese Erhöhung der Hydroxylase-Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere

10

mindestens 600 % der Hydroxylase-Aktivität des Wildtyps.

Die Bestimmung der Hydroxylase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. Referenz-

15

Organismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Aktivität der Hydroxylase wird nach Bouvier et al. (Biochim. Biophys. Acta 1391 (1998), 320-328) *in vitro* bestimmt. Es wird zu einer bestimmten Menge an Organismenextrakt Ferredoxin, Ferredoxin-NADP Oxidoreductase, Katalase, NADPH sowie beta-Carotin mit Mono- und

20

Digalaktosylglyzeriden zugegeben.

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, Keller, d'Harlingue und Camara (Xanthophyll biosynthesis: molecular and functional

25

characterization of carotenoid hydroxylases from pepper fruits (*Capsicum annuum* L.; Biochim. Biophys. Acta 1391 (1998), 320-328):

Der *in-vitro* Assay wird in einem Volumen von 0,250 ml Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7,6), 0,025

30

mg Ferredoxin von Spinat, 0,5 Einheiten Ferredoxin-NADP+ Oxidoreduktase von Spinat, 0,25 mM NADPH, 0,010 mg beta-Carotin (in

0,1 mg Tween 80 emulgiert), 0,05 mM einer Mischung von Mono- und Digalaktosylglyzeriden (1:1), 1 Einheit Katalyse, 200 Mono- und Digalaktosylglyzeriden, (1:1), 0,2 mg Rinderserumalbumin und Organismenextrakt in unterschiedlichem Volumen. Die
5 Reaktionsmischung wird 2 Stunden bei 30°C inkubiert. Die Reaktionsprodukte werden mit organischem Lösungsmittel wie THF, Aceton oder Chloroform/Methanol (2:1) extrahiert und mittels HPLC bestimmt.

10 Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, d'Harlingue und Camara (Molecular Analysis of carotenoid cyclase inhibition; Arch. Biochem. Biophys. 346(1) (1997) 53-64):

15 Der in-vitro Assay wird in einem Volumen von 250 µl Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7,6), unterschiedliche Mengen an Organismenextrakt, 20 nM Lycopin, 250 µg an chromoplastidärem Stromaprotein aus Paprika, 0,2 mM NADP⁺, 0,2 mM NADPH und 1 mM ATP. NADP/NADPH und ATP werden
20 in 10 ml Ethanol mit 1 mg Tween 80 unmittelbar vor der Zugabe zum Inkubationsmedium gelöst. Nach einer Reaktionszeit von 60 Minuten bei 30°C wird die Reaktion durch Zugabe von Chloroform/Methanol (2:1) beendet. Die in Chloroform extrahierten Reaktionsprodukte werden mittels HPLC analysiert.

25

Ein alternativer Assay mit radioaktivem Substrat ist beschrieben in Fraser und Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15).

Die Erhöhung der Hydroxylase-Aktivität kann durch verschiedene Wege
30 erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Expressions- und Proteinebene oder durch

Erhöhung der Genexpression von Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp.

Die Erhöhung der Genexpression der Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp kann ebenfalls durch verschiedene
5 Wege erfolgen, beispielsweise durch Induzierung des Hydroxylase-Gens durch Aktivatoren oder durch Einbringen von einer oder mehrerer Hydroxylase-Genkopien, also durch Einbringen mindestens einer Nukleinsäure kodierend eine Hydroxylase in denb Organismus der
10 Gattung Blakesleaa.

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure kodierend eine Hydroxylase durch Einbringen von mindestens einer Nukleinsäure kodierend eine
15 Hydroxylase in den Organismus der Gattung Blakesleaa.

Dazu kann prinzipiell jedes Hydroxylase-Gen, also jede Nukleinsäure, die eine Hydroxylase codiert, verwendet werden.

20 Bei genomischen Hydroxylase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall, dass der Wirtsorganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechende Hydroxylase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu
25 verwenden.

Ein Beispiel für ein Hydroxylase-Gen ist eine Nukleinsäure, kodierend eine Hydroxylase aus *Haematococcus pluvialis* mit der Accession No. AX038729 (WO 0061764; Nukleinsäure: SEQ ID NO: 31, Protein: SEQ ID
30 NO: 32), aus *Erwinia uredovora* 20D3 (ATCC 19321, Accession No. D90087; Nukleinsäure: SEQ ID NO: 33, Protein: SEQ ID NO: 34) oder

Hydroxylase aus *Thermus thermophilus* (DE 102 34 126.5) kodiert durch die Sequenz mit der SEQ ID NO 76.

Weitere Hydroxylasen werden von den Nukleinsäuren mit den folgenden

5 Accession Nummern kodiert

|emb|CAB55626.1, CAA70427.1, CAA70888.1, CAB55625.1,
AF499108_1, AF315289_1, AF296158_1, AAC49443.1, NP_194300.1,
NP_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1,
AAL80006.1, AF162276_1, AAO53295.1, AAN85601.1, CRTZ_ERWHE,
10 CRTZ_PANAN, BAB79605.1, CRTZ_ALCSP, CRTZ_AGRAU,
CAB56060.1, ZP_00094836.1, AAC44852.1, BAC77670.1, NP_745389.1,
NP_344225.1, NP_849490.1, ZP_00087019.1, NP_503072.1,
NP_852012.1, NP_115929.1, ZP_00013255.1

15 In den erfindungsgemäßen bevorzugten transgenen Organismen der Gattung *Blakeslea* liegt also in dieser bevorzugten Ausführungsform gegenüber dem Wildtyp mindestens ein Hydroxylase-Gen vor.

In dieser bevorzugten Ausführungsform weist der genetisch veränderte
20 Organismus beispielsweise mindestens eine exogene Nukleinsäure, kodierend eine Hydroxylase auf.

Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Ausführungsform als Hydroxylase-Gene Nukleinsäuren, die Proteine
25 kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens 50 %, bevorzugter mindestens 70%, noch
30 bevorzugter mindestens 80 %, am bevorzugtesten mindestens 90%, insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% auf

Aminosäureebene mit der Sequenz SEQ. ID. NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76 und die die enzymatische Eigenschaft einer Hydroxylase aufweisen.

- 5 Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, wie vorstehend beschrieben, durch Homologievergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der SEQ
10 ID. NO: 31, 33 oder 76 leicht auffinden.

- Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 31, 33 oder 76 aus verschiedenen Organismen deren genomische Sequenz nicht
15 bekannt ist, wie vorstehend beschrieben, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

- In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der Hydroxylase-Aktivität Nukleinsäuren in Organismen
20 eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der Hydroxylase der Sequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76.

- Geeignete Nukleinsäuresequenzen sind beispielsweise durch
25 Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

- Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Organismenspezifischen codon usage häufig verwendet werden. Die
30 codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 31, 33 oder 76 in den Organismus ein.

5

Alle vorstehend erwähnten Hydroxylase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Der in der Transformation (i) eingesetzte Vector umfasst daher in weiteren Ausführungsformen der Erfindung bevorzugterweise eine Sequenz codierend für eine Hydroxylase, insbesondere eine Hydroxylase aus *Haematococcus pluvialis* mit der SEQ ID NO: 70 oder eine Hydroxylase aus *Erwinia uredovae* mit der SEQ ID NO: 71. oder eine Hydroxylase aus *Thermus thermophilus* kodiert durch die Sequenz mit der SEQ ID NO 76.

25

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise ferner die Expression regelnde und unterstützende Bereiche, insbesondere Promotoren und Terminatoren.

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise den *gpd* und/oder den *ptef1* Promotor und/oder den *trpC* Terminator. Diese

30

haben sich zur Transformation der *Blakeslea* besonders bewährt. Auch der Einsatz von dem Fachmann geläufigen "inverted repeats" (IR, Römpf Lexikon der Biotechnologie 1992, Thieme Verlag Stuttgart, Seite 407 "Invers repetitive Sequenzen") zur Regelung der Expression bzw.
5 Transkription liegt im Rahmen der Erfindung.

Vorteilhafterweise weist der im Vector eingesetzte *gpd* Promotor die Sequenz SEQ ID NO: 1 auf. Vorteilhafterweise weist der im Vector eingesetzte *trpC* Terminator die Sequenz SEQ ID NO: 2 auf.
10 Vorteilhafterweise weist der im Vector eingesetzte *ptef1* Promotor die Sequenz SEQ ID NO: 35 auf.

Insbesondere werden dabei der *gpd* Promotor und der *trpC* Terminator aus *Aspergillus nidulans* und der *ptef1* Promotor aus *Blakeslea trispora*
15 eingesetzt.

Insbesondere enthält der in der Transformation (i) eingesetzte Vector ein Resistenzgen. Bevorzugterweise handelt es sich um ein Hygromycin-Resistenzgen (*hph*), insbesondere das aus *E. coli*. Dieses Resistenzgen
20 hat sich bei dem Nachweis der Transformation und Selektion der Zellen als besonders geeignet herausgestellt.

Als Promotor für *hph* wird also bevorzugt *p-gpdA*, der Promotor der Glycerinaldehyd-3-phosphatdehydrogenase aus *Aspergillus nidulans*
25 genutzt. Als Terminator für *hph* wird bevorzugt *t-trpC*, der Terminator des Gens *trpC*, codierend für Anthranilatsynthasekomponenten aus *Aspergillus nidulans* genutzt.

Als Vektoren haben sich Abkömmlinge des *pBinAHyg* Vectors als
30 besonders geeignet herausgestellt. Der zur Transformation eingesetzte Vector umfasst also bevorzugterweise die SEQ ID NO: 3.

Hinzu kommen je nach gewünschtem Carotinoid oder dessen Vorstufe eine Sequenz codierend für eine Hydroxylase, Ketolase, Phytoendesaturase usw. wie diese zuvor beschrieben wurden. Die Vektoren umfassen also in einer Ausführungsform der Erfindung die

5 Sequenz SEQ ID NO: 69 codierend für die Phytoendesaturase. Die Vektoren umfassen ferner in einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 72 codierend für eine Ketolase. Die Vektoren umfassen weiter in einer weiteren Ausführungsform der Erfindung die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase.

10 Entsprechende Kombinationen der zuvorgenannten Sequenzen liegen ebenso im Rahmen der Erfindung. So umfasst der Vector in einer Ausführungsform sowohl eine Sequenz SEQ ID NO: 72 codierend für eine Ketolase als auch die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydroxylase und ermöglicht so die Herstellung von Astaxanthin.

15 Insbesondere sind Vektoren ausgewählt aus der Gruppe bestehend aus den SEQ ID NO: 37 bis 51 und 62 im Rahmen der Erfindung einsetzbar.

Mit dem erfindungsgemäßen Verfahren sind gentechnisch veränderte

20 Organismen *Blakeslea*, insbesondere der Art *Blakeslea trispora* bzw. aus ihnen gebildetes Mycel erhältlich.

Die genetisch veränderten Organismen können zur Produktion von Carotinoiden, Xanthophyllen oder deren Vorstufen, insbesondere Phytoen,

25 Bixion, Astaxanthin, Zeaxanthin und Canthaxanthin verwendet werden. Auch können neue, im Wildtyp natürlicherweise nicht vorkommende Carotinoide durch Einbringung der entsprechenden genetischen Information von den gezielt genetisch veränderten Zellen bzw. dem durch sie gebildeten Mycel erzeugt und anschließend isoliert werden.

Bevorzugterweise ist die Gewinnung von Carotinoiden oder deren Vorstufen mit den gezielt genetisch veränderten Zellen bzw. das durch sie gebildete Mycel möglich.

- 5 Wird die gentechnische Veränderung nur in Zellen eines der vorkommenden Paarungstypen (bei *Blakeslea trispora* (+) oder (-)) durchgeführt, so wird zur Kultivierung der entsprechend andere, nicht veränderte Paarungstyp zugesetzt, da so eine gute Produktion der Carotinoide oder deren Vorstufen aufgrund der von dem zweiten, nicht
- 10 veränderten Paarungstyp abgegebenen Substanzen (z. B. Trisporsäuren) zu erreichen ist. Vorteilhafterweise wird jedoch die gentechnische Veränderung in Zellen beider Paarungstypen vorgenommen und diese zusammen kultiviert. Hierdurch wird ein besonders gutes Wachstum und eine optimale Produktion der Carotinoiden oder deren Vorstufen erreicht.
- 15 Auch eine (künstliche) Zugabe der Trisporsäuren ist möglich und sinnvoll.

- Trisporsäuren sind Sexualhormone in Mucorales Pilzen, wie *Blakeslea*, welche die Bildung von Zygothoren und die Produktion von β -Carotin stimulieren (van den Ende 1968, J. Bacteriol. 96:1298 - 1303, Austin et al.
- 20 1969, Nature 223:1178 – 1179, Reschke Tetrahedron Lett. 29:3435 – 3439, van den Ende 1970, J. Bacteriol. 101:423 – 428).

Die Erfindung wird nachfolgend an Hand von Beispielen näher ausgeführt.

25 **Material und Methoden**

Molekulargenetische Arbeiten wurden, wenn nicht anders beschrieben, nach den Methoden in Current Protocols in Molecular Biology (Ausubel et al., 1999, John Wiley & Sons) durchgeführt.

30 **Stämme und Wachstumsbedingungen**

Die *Blakeslea trispora* Stämme ATCC 14271 (Paarungstyp(+)) und ATCC14272 (Paarungstyp (-)) wurden von der American Type Culture Collection erhalten. Die Anzucht von *B. trispora* erfolgte in MEP-Medium (Malzextrakt-Pepton-Medium): 30 g/l Malzextrakt (Difco), 3 g/l Pepton
5 (Soytone, Difco), 20 g/l Agar, Einstellung pH 5,5, ad 1000 ml mit H₂O bei 28 °C.

Die Anzucht von *Agrobacterium tumefaciens* LBA4404 erfolgte nach Hoekema et al. (1983, Nature 303:179-180) bei 28 °C für 24 h in
10 *Agrobacterien*-Minimal Medium (AMM): 10 mM K₂HPO₄, 10 mM KH₂PO₄, 10 mM Glucose, MM-Salze (2,5 mM NaCl, 2 mM MgSO₄, 700 µM CaCl₂, 9 µM FeSO₄, 4 mM (NH₄)₂SO₄).

Transformation von *Agrobacterium tumefaciens*

15 Das Plasmid pBinAHyg wurde in den *Agrobakterien*stamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der *Agrobakterien*anzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l
20 (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Transformation von *Blakeslea trispora*

Zur Transformation wurden die *Agrobakterien* nach 24 h Anzucht in AMM
25 auf eine OD₆₀₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₀₀ von ca. 0,6 angezogen.

30 Zur Co-Inkubation von *Blakeslea* ATCC 14271 bzw. ATCC14272 und *Agrobacterium* wurden 100 µl *Agrobakterien*suspension mit 100 µl

Blakeslea Sporensuspension (10^7 Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l
5 Pepton, pH 5,5, 18 g/l Agar) überführt. Zur Selektion auf transformierte Blakesleazellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete
10 Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l $MgSO_4 \times 7H_2O$, 150 mg/l KH_2PO_4 , 25 µg/l ThiaminHCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, 100 mg/L Hygromycin, 100 mg/L Cefotaxim, pH 5,5, 18 g/l Agar) ausplattiert. Zur Isolierung einzelner gentechnisch veränderter Sporen
15 wurden die Sporen durch ein FACS Gerät der Fa. BectonDickson (Modell Vantage+Diva Option) einzeln auf Selektivmedium abgelegt.

Herstellung genetisch veränderter Blakeslea trispora durch Agrobacterium-vermittelte Transformation

20 Herstellung des rekombinanten Plasmids pBinAHyg

Aus dem Plasmid pANsCos1 (Fig.1, Osiewacz, 1994, Curr. Genet. 26:87-90, SEQ ID NO: 4) wurde die gpdA-hph-trpC-Kassette als BglII/HindIII Fragment isoliert und in das mit BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so
25 erhaltene Vektor wurde als pBinAHyg bezeichnet (Fig. 2, SEQ ID NO: 3) und enthielt das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors (SEQ ID NO: 1) und des trpC Terminators (SEQ ID NO: 2) aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium*
30 notwendig sind. Die in den weiter unten beschriebenen

Ausführungsbeispielen genannten Vektoren sind Abkömmlinge von pBinAHyg.

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in

5 Agrobacterium tumefaciens

Nachfolgend wird beispielhaft die Übertragung des Plasmids pBinAHyg in Agrobakterien beschrieben. Die Übertragung der Abkömmlinge erfolgte analog.

- 10 Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das *A. tumefaciens* Chromosom), Streptomycin 30 mg/l
15 (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in Blakeslea trispora

- 20 Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD₆₆₀ von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 µM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₆₀ von ca. 0,6 angezogen.

25

- Zur Co-Inkubation von *Blakeslea trispora* (B.t.) und *Agrobacterium tumefaciens* (A.t.) wurden 100 µl Agrobaktériensuspension mit 100 µl Blakeslea Sporensuspension (10⁷ Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose
30 Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C

wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt.

Zur Selektion auf transformierte Blakeslea-Zellen enthielt das Medium
5 Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion
gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7
Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische
Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und
auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l $\text{MgSO}_4 \times$
10 $7\text{H}_2\text{O}$, 150 mg/l KH_2PO_4 , 25 $\mu\text{g/l}$ Thiamin-HCl, 100 mg/l Yeast Extract,
100 mg/l Na-desoxycholat, pH 5,5, 100 mg/l Cefotaxim, 100 mg/l
Hygromycin, 18 g/l Agar) ausplattiert. Die Übertragung von Sporen auf
frische Selektionsplatten wurde dreimal wiederholt. Auf diese Weise wurde
die Transformante *Blakeslea trispora* GVO 3005 isoliert. Alternativ erfolgte
15 zur Selektion der GVO (gentechnisch veränderten Organismen) die
Einzelablage der Sporen durch den Becton Dickinson FacsVantage+Diva
Option auf CM-17 Agar mit 100 mg/l Cefotaxim, 100 mg/l Hygromycin. In
diesem Fall wurde nur dort Pilzmycel gebildet, wo die Sporen
gentechnisch verändert waren.

20

**Nachweis der genetischen Veränderung durch Übertragung von
pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora***

Nachfolgend wird beispielhaft der Nachweis der Übertragung für
pBinAHyg in *Blakeslea trispora* beschrieben. Der Nachweis der
25 Übertragung der Abkömmlinge erfolgte analog.

200 ml MEP-Medium (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5) wurden mit
 10^5 bis 10^7 Sporen der Transformante *Blakeslea trispora* GVO 3005
beimpft und 7 Tage bei 26 °C mit 200 Upm auf einem Rundschüttler
30 inkubiert. Zum Nachweis der erfolgreichen Transformation wurde DNA aus
dem Mycel isoliert (Pqclab Fungal DNA Mini Kit) und in einer PCR

(Programm: 94 °C 1 min, dann 30 Zyklen mit 1 min. 94°C, 1 min. 58 °C, 1 min. 72 °C) eingesetzt.

Zum Nachweis des Hygromycinresistenzgens (hph) wurden die Primer
5 hph-forward (5'-CGATGTAGGAGGGCGTGGATA, SEQ ID NO: 5) und
hph-reverse (5'-GCTTCTGCGGGCGATTTGTGT, SEQ ID NO: 6)
verwendet. Das erwartete Fragment von hph wies eine Länge von 800 bp
auf.

10 Zur Amplifikation des Kanamycinresistenzgens nptIII und damit als
Kontrolle auf Agrobakterien wurden die Primer nptIII-forward (5'-
TGAGAATATCACCGGAATTG, SEQ ID NO: 7) und nptIII-reverse (5'-
AGCTCGACATACTGTTCTTCC, . SEQ ID NO: 8) verwendet. Das
erwartete Fragment von nptIII wies eine Länge von 700 bp auf.

15

Zur Amplifikation eines Fragmentes des Glycerinaldehyd-3-
phosphatdehydrogenasegens gpd1 und damit als Kontrolle auf Blakeslea
trispota wurden die Primer MAT292 (5'-
GTGAATGGAAATCCCATCGCTGTC, SEQ ID NO: 9) und MAT293 (5'-
20 AGTGGGTACTCTAAAGGCCATACC, SEQ ID NO: 10) verwendet. Das
erwartete Fragment von gpd1 wies eine Länge von 500 bp auf.

Das Ergebnis der PCR der Blakeslea trispota DNA ist in Fig. 3 anhand
eines Standard-Gels gezeigt. Die Spuren des Gels wurden
25 folgendermaßen belegt:

- 1) 100 bp Größenmarker (100 bp - 1 kb)
- 2) B.t. GVO 3005 primer nptIII-for / nptIII-rev
- 3) B.t. GVO 3005 primer hph-for / hph-rev
- 30 4) B.t. GVO 3005 primer MAT292 / MAT293 (gpd)
- 5) A.t. mit Plasmid pBinAHyg primer nptIII-for / nptIII-rev

- | | |
|------------------------------|--------------------------------|
| 6) A.t. mit Plasmid pBinAHyg | primer hph-for / hph-rev |
| 7) B.t. 14272 WT | primer nptIII-for / nptIII-rev |
| 8) B.t. 14272 WT | primer hph-for / hph-rev |
| 9) B.t. 14272 WT | primer MAT292 / MAT293 (gpd) |

5

In der DNA von *Blakeslea trispora* wurde das Hygromycinresistenzgens (hph) und als Positivkontrolle Glycerinaldehyd-3-phosphatdehydrogenasegen (gpd1) nachgewiesen. nptIII konnte demgegenüber nicht nachgewiesen werden.

10

Somit wurde die genetische Veränderung von *Blakeslea trispora* durch *Agrobacterium*-vermittelte Transformation nachgewiesen.

Isolierung homokaryotischer GVO von *Blakeslea trispora*:

- 15 Durch erfolgreichen Transfer des Vectors pBinAHyg und Abkömmlingen von pBinAHyg in *Blakeslea trispora* entstehen genetisch veränderte Organismen (GVO) von *Blakeslea trispora*. Jedoch liegen in *Blakeslea* in allen Stadien des vegetativen und des sexuellen Zellzyklus mehrkernige Zellen vor. Daher erfolgt die Insertion der Fremd-DNA in der Regel nur in
- 20 einem Kern. Ziel ist es, Stämme von *Blakeslea* zu erhalten, bei denen die Insertion der Fremd-DNA in allen Kernen vorliegt, d.h. Ziel ist ein homonukleates rekombinantes Pilzmycel.

- 1) Herstellung homonukleater rekombinanter Stämme durch FACS**
- 25 **(fluorescence-activated cell sorting)**

- Ein geringer Anteil der Sporen von *Blakeslea trispora* bzw. der gentechnisch veränderten Stämme von *Blakeslea trispora* ist von Natur aus einkernig. Zur Herstellung homonukleater rekombinanter Stämme, die Fremd-DNA von pBinAHyg oder pBinAHyg-Abkömmlingen enthielten,
- 30 wurden die einkernigen Sporen durch FACS aussortiert und auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim

und 100 mg/l Hygromycin plattiert. Die hier gebildeten Mycelien waren homonukleat. Zur Sortierung mit FACS wurden die Sporen eines 3 Tage alten Ausstriches mit 10 ml Tris-HCl 50mMol + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis $0,8 \times 10^7$ Sporen pro ml. Zu 9 ml Sporensuspension wurden 1ml DMSO und 10 µl Syto 11 (Farbstoff-Stammlösung in DMSO Molecular Probes Nr.S-7573) zugegeben. Danach wurde 2 h bei 30°C gefärbt. Selektion und Ablage erfolgten mittels eines Gerätes vom Typ FacsVantage+Diva Option Fa. Becton Dickinson. Die Selektion erfolgte zuerst nach Größe, um einzelne Sporen von Aggregaten und Verunreinigungen zu trennen. Dann wurden diese Sporen nach ihrer Fluoreszenz (Anregung = 488nm; Emission = 530 nm) sortiert abgelegt. Die linke Schulter der Gauß-Kurve der Fluoreszenzhäufigkeitsverteilung enthielt die einkernigen Sporen.

2) Herstellung homonukleater Stämme durch Kernreduktion und Selektion mit FACS

Zur Reduzierung der Anzahl von Kernen pro Spore wurde vor der Selektion eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt, und so durch chemische Mutagenese eine Kernreduktion erzielt.

Hierfür wurde zunächst eine Sporensuspension mit 1×10^7 Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und nach der unter 1) beschriebenen Methode sortiert bzw. selektiert.

Alternativ konnten zur Reduktion der Kernzahl in den Sporen auch Röntgen – und UV-Strahlen eingesetzt werden, wie es von Cerdá-Olmedo und Patricia Reau in Mutation Res., 9 (1970), 369-384 beschrieben wurde.

5 **3) Herstellung homonukleater Stämme durch Selektion auf rezessive Selektionsmarker**

Als rezessiver Selektionsmarker zur Selektion homonukleater Mycelien kommt beispielsweise der rezessive Selektionsmarker pyrG in Frage. Wildtyp-Stämme von *Blakeslea trispora* sind pyrG⁺. Diese Stämme können
10 nicht in Gegenwart des Pyrimidin-Analogs 5-Fluororotat (FOA) wachsen, weil sie FOA durch die Orotidin-5'-monophosphatdecarboxylase zu lethalen Metaboliten umsetzen. Gentechnisch veränderte *Blakesleaa*, die homonukleat pyrG⁻ sind, fehlt die Enzymaktivität Orotidin-5'-monophosphatdecarboxylase. Folglich können diese pyrG⁻-Stämme 5-
15 Fluororotat nicht verwerten. Die Stämme wachsen daher in Gegenwart von FOA und Uracil. Im Fall der Kopplung der Mutation pyrG⁻ und der Insertion von Fremd-DNA auf dem Kern einer einkernigen Spore, kann aus dieser Spore homonukleates rekombinantes Pilzmycel gebildet werden.

20

Zunächst wurde durch Insertion eines Fragmentes von pyrG (SEQ ID NO: 65) aus *Blakeslea trispora* in pBinAHyg das Plasmid pBinAHygBTpyrG-SCO (SEQ ID NO: 36, Fig. 4) erzeugt. Dieses Plasmid wurde in *Blakeslea trispora* transformiert und führte dort durch homologe Rekombination zur
25 Disruption von pyrG.

Homonukleate GVO von *Blakeslea trispora* mit dem Phänotyp pyrG⁻ wurden folgendermaßen selektiert. Zur Agrobakterium-vermittelten Transformation von pBinAHygBTpyrG-SCO wurde wie oben beschrieben
30 auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die Sporen der

Transformanten wurden mit 10 ml Tris-HCl 50mM + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis $0,8 \times 10^7$ Sporen pro ml. Die Sporen wurden anschließend auf FOA-Medium mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin ausplattiert. FOA-Medium
5 enthielt pro Liter 20 g Glucose, 1 g FOA, 50 mg Uracil, 200 ml Citrat-Puffer (0,5 M, pH 4,5) und 40 ml Spurensalzlösung nach Sutter, 1975, PNAS, 72:127). Homonukleate pyrG⁻-Mutanten zeigten Wachstum auf dem Uracil-haltigen FOA-Medium; aber kein Wachstum bei Plattierung auf FOA-Medium ohne Uracil. Auf die gleiche Weise wurden aus den im
10 folgenden beschriebenen GVO von *Blakeslea trispora* zur Herstellung von Xanthophyllen homonukleate GVO hergestellt.

Alternativ ist es möglich die Sporen analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin zu plattieren, das
15 zusätzlich Hygromycin enthält (Roncero et al., 1984, Mutation Research, 125: 195 - 204). Hierdurch werden homokaryonte Zellen des Genotyps hyg^R und dar⁻ selektiert. Nach diesem Prinzip werden homokaryonte Stämme von *Blakeslea trispora* mit dem Phänotyp hyg^R und dar⁻ erzeugt.

20

Ausführungsbeispiele zur Herstellung von gentechnisch veränderten Organismen von *Blakeslea trispora* für die Herstellung von Carotinoiden und Carotinoidvorstufen

25 Die Erzeugung der im folgenden genannten Plasmide erfolgte durch die Methode „overlap-extension PCR“ und durch anschließende Insertion der Amplifikationsprodukte in das Plasmid pBinAHyg. Die Methode „overlap-extension PCR“ erfolgte wie in Innis et al. (Eds.) PCR protocols: a guide to methods and applications, Academic Press, San Diego beschrieben. Die
30 Transformation der pBinAHyg-Abkömmlinge und die Herstellung

homonukleater gentechnisch veränderter Stämme von *Blakeslea trispora* erfolgte wie oben beschrieben.

Gentechnisch veränderte Stämme von *Blakeslea trispora* zur

5 Herstellung von Zeaxanthin

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Zeaxanthin verwendet, codieren also u.a. Hydroxylasen (crtZ):

- 10 - p-tef1-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ (SEQ ID NO: 70) aus *Haematococcus pluvialis* Flotow NIES-144 (Accession No. AF162276) unter Kontrolle des ptef1 Promotors aus *Blakeslea trispora* (Seq. pBinAHygBTpTEF1-HPcrtZ, SEQ ID NO: 37, Fig. 5);
- 15 - p-carRA-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHyg-BTpcarRA-HPcrtZ, SEQ ID NO: 38, Fig. 6)
- p-carB-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-HPcrtZ, SEQ ID NO: 39, Fig. 7)
- 20 - p-carRA-HPcrtZ-TAG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Stromabwärts des Gens der Hydroxylase ist eine Inverted-Repeat-Struktur lokalisiert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1' ca. 350 bp von carA, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2') (Seq. pBinAHyg-BTpcarRA-HPcrtZ-TAG-3'carA-IR, SEQ ID NO: 40, Fig. 8);
- 25

- p-carRA-HPcrtZ-GCG-3'carA-IR, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora*. Das Gen der Hydroxylase ist mit einer Inverted-Repeat-Struktur fusioniert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, 'Inverted Repeat 1' ca. 350 bp von carA, dann ca. 200 bp 'Loop' und anschließend ca. 350 bp 'Inverted Repeat 2'). Das abgeleitete Fusionsprotein besteht folglich aus der Hydroxylase von *Haematococcus pluvialis* und dem Carboxyterminus von CarA aus *Blakeslea trispora* (Seq. pBinAHyg-BTpcarRA-HPcrtZ-GCG-3'carA-IR, SEQ ID NO: 41, Fig. 9);
- p-tef1-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ (SEQ ID NO: 71) aus *Erwinia uredova* 20D3 (Accession No. D90087) unter Kontrolle des ptef1 Promotors (Seq. pBinAHygBTpTEF1-EUcrtZ, SEQ ID NO: 42, Fig. 10);
- p-carRA-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus *Erwinia uredova* 20D3 unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-EUcrtZ, SEQ ID NO: 43, Fig. 11);
- p-carB-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus *Erwinia uredova* 20D3 unter Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-EUcrtZ, SEQ ID NO: 44, Fig. 12);
- p-gpdA-HPcrtZ-t-crtZ, enthaltend Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 unter Kontrolle des gpdA Promotors und des Terminators t-crtZ; d.h. des stromabwärts von crtZ aus *Haematococcus pluvialis* Flotow NIES-144 gelegenen Sequenzabschnitts (SEQ ID NO: 73) (Seq. pBinAHyg-gpdA-HPcrtZ-tcrtZ, SEQ ID NO: 45, Fig. 13).

- 5 - p-gpdA-BTcarR-HPcrtZ-BTcarA, enthaltend Genfusion aus Genen der Lycopincyclase carR aus *Blakeslea trispora*, der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 und der Phytoensynthase carA aus *Blakeslea trispora* unter Kontrolle des gpdA Promotors aus *Aspergillus nidulans* (Seq. pBinAHyg-carR_crtZ_carA, SEQ ID NO: 46, Fig. 14);

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Canthaxanthin

10 Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung von Canthaxanthin verwendet, codieren also u.a. Ketolasen (crtW):

- 15 - p-tef1-NPcrtW, enthaltend das Gen der Ketolase NPcrtW (SEQ ID NO: 72) aus *Nostoc punctiforme* PCC73102 (ORF148, Accesion No. NZ_AABC01000196) unter Kontrolle des ptef1 Promotors aus *Blakeslea trispora* (Seq. pBinAHygBTpTEF1-NpucrtW, SEQ ID NO: 47, Fig. 15);
- 20 - p-carRA-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-NpucrtW, SEQ ID NO: 48, Fig. 16);
- 25 - p-carB-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 unter der Kontrolle des Promotors pcarB aus *Blakeslea trispora* (Seq. pBinAHygBTpcarB-NpucrtW, SEQ ID NO: 49, Fig. 17);

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Astaxanthin

30 Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von *Blakeslea trispora* für die Herstellung

von Astaxanthin verwendet, codieren also u.a. für Hydroxylasen (crtZ) und Ketolasen (crtW):

- p-carRA-HPcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase HPcrtZ aus *Haematococcus pluvialis* Flotow NIES-144 und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 (ORF148, Accession No. NZ_AABC01000196) beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 50, Fig. 18);
- p-carRA-EUcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase EUcrtZ aus *Erwinia uredovae* 20D3 (Accession No. D90087) und das Gen der Ketolase NPcrtW aus *Nostoc punctiforme* PCC73102 beide jeweils unter Kontrolle des Promotors pcarRA aus *Blakeslea trispora* (Seq. pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 51, Fig. 19);

Klonierung und Sequenzanalyse von Genen und Promotoren, die beispielhaft für die gentechnische Veränderung von *Blakeslea trispora* genutzt werden können.

- Nachfolgend werden beispielhaft die Klonierung und Sequenzierung verschiedener Gene und Promotoren aus *Blakeslea trispora* beschrieben.

Klonierung und Sequenzanalyse ptef1

- Die Klonierung von p-tef aus *Blakeslea trispora* erfolgte auf der Grundlage einer bereits in GenBank veröffentlichten Sequenz des Strukturgens für den Translations-Elongationsfaktor 1- α aus *Blakeslea trispora* (AF157235). Ausgehend von dem Sequenzeintrag AF157235 wurden Primer für die inverse PCR ausgewählt, um die stromaufwärts des Strukturgens gelegene Promotoregion zu amplifizieren und zu sequenzieren.

In der inversen nested PCR an 200 ng XhoI-gespaltener und zirkularisierter genomischer DNA von *Blakeslea trispora* ATCC14272 wurde ein 3000-bp-Fragment in folgendem Ansatz erhalten: Matrizen-DNA (1 µg genomische DNA von *Blakeslea trispora* ATCC 14272) Primer
5 MAT344 5'-GGCGTACTTGAAGGAACCCCTTACCG-3' (SEQ ID NO: 63)
und MAT 345 5'-ATTGATGCTCCCGGTCACCGTGATT-3' (SEQ ID NO: 64) je 0,25 µM, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase (Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C,
10 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus). Der Sequenzabschnitt, der stromaufwärts des vermutlichen Startcodons des Gens *tef1* innerhalb 3000-bp-Fragmentes liegt, wurde als Promotor *ptef1* bezeichnet.

Klonierung Sequenzanalyse des Gens der HMG-CoA-Reduktase aus 15 *Blakeslea trispora*

Zunächst wurde mit dem Cosmidvektor pANsCos1 eine Genbank von *Blakeslea trispora* ATCC 14272, Mating Type (–) hergestellt. Der Vektor wurde durch Spaltung mit XbaI linearisiert und anschließend dephosphoryliert. Eine weitere Spaltung mit BamHI schuf die
20 Insertionsstelle, in welche die mit Sau3AI partiell gespaltene und dephosphorylierte genomische DNA von *Blakeslea trispora* ligiert wurde. Die derart gebildeten Cosmide wurden anschließend *in vitro* verpackt und in *Escherichia coli* übertragen.

Auf der Grundlage der bekannten Sequenz eines Fragmentes des HMG-CoA-Reduktase codierenden Gens aus *Blakeslea trispora* (Eur. J. Biochem 220, 403–408 (1994)) wurde eine 315-bp-DNA-Sonde durch folgende PCR hergestellt. Reaktionsansatz: 1 µg genomische DNA von *Blakeslea trispora* ATCC 14272, Primer MAT314 5'-
25 CCGATGGCGACGACGGAAGGTTGTT-3' [SEQ ID NO 79] und MAT315
30 5'-CATGTTTCATGCCCATTGCATCACCT-3' [SEQ ID NO 80] je 0,25 µM, 100 µM dNTP, 10 µl Herculanase-Polymerasepuffer 10x, 5 U Herculanase

(Zugabe bei 85 °C), H₂O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 58 °C, 30 s. 72 °C, 30 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus).

5 Mit dieser DNA-Sonde wurde die Cosmid-Genbank durchmustert. Es wurde ein Klon identifiziert, dessen Cosmid mit der DNA-Sonde hybridisierte. Die Insertion dieses Cosmids wurde sequenziert. Die DNA-Sequenz enthielt einen Abschnitt, der dem Gen einer HMG-CoA-Reduktase zugeordnet wurde [SEQ ID NO 75].

10 Klonierung und Sequenzanalyse carB

(carB = Gen der Phytoendesaturase aus *Blakeslea trispora*)

Aus dem Sequenzvergleich der Peptidsequenzen von Phytoendesaturasen und dem Vergleich der zugehörigen DNA-Sequenzen von *Phycomyces blakesleeanus*, *Cercospora nicotianae*,
15 *Phaffia rhodozyma* und *Neurospora crassa* wurden die degenerierten Primer MAT182 5'-GCNGARGGNATHHTGGTA-3' (SEQ ID 52) und MAT192 5'-TCNGCNAGRAADATRTTTRTG-3' (SEQ ID 53) abgeleitet. Die PCR wurde in 100 µl Ansätzen durchgeführt. Diese enthielten 200 ng genomische DNA von *Blakeslea trispora* ATCC14272, 1 µM MAT182, 1
20 µM MAT192, 100 µM dNTP, 10 µl Pfu-Polymerasepuffer 10x, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C), H₂O ad 100 µl.

Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 40 °C, 30 s, 72 °C, 30 s, 95 °C, 30 s (35 Zyklen); 72 °C, 10 min (1 Zyklus).

25

Hiermit wurde ein 358-bp-Fragment erhalten, dessen abgeleitete Peptidsequenz Ähnlichkeit zu den Sequenzen der Phytoendesaturasen aufwies. Durch die Methode der inversen PCR (Innis et al. in PCR protocols: a guide to methods and applications. 1990. S. 219-227) wurden
30 nach dem Prinzip des Chromosome-Walking die Genregionen

stromaufwärts und stromabwärts des 350-bp-Fragmentes folgendermaßen amplifiziert, kloniert und sequenziert:

- (i) ein 1,1-kbp-Fragment durch PCR mit den Primern MAT219 5'-AAGTGACACCGGTTACACGCTTGTCTT-3' (SEQ ID 54) und MAT
5 220 5'-GCTTATCACCATCTGTTACCTCCTTGC-3' (SEQ ID 55)
erhalten aus 200 ng EcoRI-gespaltener und zirkularisierter
genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM
MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculase-
Polymerasepuffer 10x, 5 U Herculase (Zugabe bei 85 °C), H₂O ad
10 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min
(1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C,
10 min (1 Zyklus),
- (ii) ein 2,9-kbp-Fragment durch PCR mit den Primern MAT219 und
MAT220 erhalten aus 200 ng XbaI-gespaltener und zirkularisierter
15 genomischer DNA von *Blakeslea trispora* ATCC14272, 0,25 µM
MAT219, 0,25 µM MAT220, 100 µM dNTP, 10 µl Herculase-
Polymerasepuffer 10x, 5 U Herculase (Zugabe bei 85 °C), H₂O ad
100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min
(1 Zyklus); 60 °C, 30 s, 72 °C, 3 min, 95 °C, 30 s (30 Zyklen); 72 °C,
20 10 min (1 Zyklus);

Der klonierte Sequenzabschnitt ist schematisch in Fig. 20 [SEQ ID NO 77]
dargestellt. Die Sequenzierung erfolgte in Strang- und
Gegenstrangrichtung mit den klonierten Fragmenten sowie mit den PCR-
Produkten. Die Sequenz des klonierten Sequenzabschnitts ist in Fig. 21
25 [SEQ ID NO 78] gezeigt.

Sequenzvergleiche

Die Nukleotidsequenz von *carB* und die Peptidsequenz des abgeleiteten
Proteins CarB wurden mit den bekannten Sequenzen verwandter Proteine
30 verglichen. Zum Sequenzvergleich wurden die Programme GAP und
BESTFIT eingesetzt.

CarB - Identische Aminoacylreste nach GAP

Programmeinstellungen:

Gap Weight: 8

5 Length Weight: 2

Average Match: 2.912

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

10 *Phycomyces blakesleeanus*: 72,491

Phaffia rhodozyma: 50,460

Neurospora crassa: 47,943

Cercospora nicotianae: 47,740

15

CarB - Identische Aminoacylreste nach BESTFIT

Programmeinstellungen:

Gap Weight: 8

Length Weight: 2

20 Average Match: 2.912

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu CarB aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 73,380

25 *Phaffia rhodozyma*: 53,175

Neurospora crassa: 51,896

Cercospora nicotianae: 50,791

carB - Identische Basen nach GAP

30 Programmeinstellungen:

Gap Weight: 50

Length Weight: 3
Average Match: 10.000
Average Mismatch: 0.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB
5 aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 64,853
Cercospora nicotianae: 50,143
Phaffia rhodozyma: 43,179
Neurospora crassa: 42,130

10

carB -Identische Basen nach BESTFIT

Programmeinstellungen:

Gap Weight: 50
Length Weight: 3
15 Average Match: 10.000
Average Mismatch: -9.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB
aus *Blakeslea trispora* ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 68,926
20 *Phaffia rhodozyma*: 62,403
Neurospora crassa: 60,230
Cercospora nicotianae: 56,884

Klonierung zur Expression von carB

25 Zur Klonierung und Expression von carB aus *Blakeslea trispora* wurden
von dem oben beschriebenen klonierten Sequenzabschnitt aus *Blakeslea*
trispora in sechs Leserastern die möglichen Proteinsequenzen abgeleitet.
Diese Proteinsequenzen wurden mit den Sequenzen der
Phytoendesaturasen aus *Phycomyces blakesleeanus*, *Phaffia rhodozyma*,
30 *Neurospora crassa*, *Cercospora nicotianae* verglichen. Auf der Grundlage
des Sequenzvergleiches wurden im klonierten Sequenzabschnitt der

genomischen DNA von *Blakeslea trispora* drei Exons identifiziert, die zusammengefügt eine codierende Region ergeben, deren abgeleitetes Genprodukt über die gesamte Länge 72,7% identische Aminoacylreste mit der Phytoendesaturase CarB aus *Phycomyces blakesleeanus* aufweist.

5 Dieser Sequenzabschnitt aus drei möglichen Exons und zwei möglichen Introns wurde daher als Gen *carB* bezeichnet. Zur Überprüfung der vorhergesagten Genstruktur wurde die codierende Sequenz von *carB* aus *Blakeslea trispora* durch PCR mit cDNA von *Blakeslea trispora* als Matrize und mit den Primern Bol1425 5'-

10 AGAGAGGGATCCTTAAATGCGAATATCGTTGC-3' (SEQ ID 56) und Bol1426 5'-AGAGAGGGATCCATGTCTGATCAAAAGAAGCA-3' (SEQ ID 57) erzeugt. Das erhaltene DNA-Fragment wurde sequenziert. Die Lokalisation von Exons und Introns wurde durch Vergleich der cDNA mit der genomischen DNA von *carB* bestätigt. In Fig. 21 ist die codierende

15 Sequenz von *carB* schematisch dargestellt. Zur Expression von *carB* in *Escherichia coli* wurde zunächst die NdeI-Schnittstelle in *carB* durch die Methode overlap extension PCR entfernt sowie am 5'-Ende des Gens eine NdeI-Schnittstelle und am 3'-Ende eine BamHI-Schnittstelle eingefügt. Das erhaltene DNA-Fragment wurde mit dem Vektor pJOE2702 ligiert.

20 Das erhaltene Plasmid wurde als pBT4 bezeichnet und zusammen mit pCAR-AE in *Escherichia coli* XL1-Blue kloniert. Die Expression erfolgte durch Induktion mit Rhamnose. Der Nachweis der Enzymaktivität erfolgte durch Nachweis der Lycopinsynthese via HPLC. Die Klonierungsschritte sind im folgenden beschrieben:

25 **PCR 1.1:**

Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT350 5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID 58),

0,25 µM MAT244 5'-GTTCCAATTGGCCACATGAAGAGTAAGACAGGAAACAG-3' (SEQ ID

30 59), 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100µL.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30 s, 9. 72 °C 10min
Zyklen: (1-2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

5

PCR1.2:

Ca. 0,5 µg cDNA von *Blakeslea trispora*, 0,25 µM MAT243 5'-CCTGTCTTACTCTTCATGTGGCCAATTGGAACCAACAC-3' (SEQ ID 60),
0,25 µM MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID 61),
10 100 µM dNTP, 10 µl Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase
(Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30s, 9. 72 °C 10min
15 Zyklen: (1 -2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

Reinigung der PCR-Fragmente aus PCR 1.1, 1.2

Dazu wurde PCR 2 zur Herstellung der codierenden Sequenz von carB
20 aus *Blakeslea trispora* für die Klonierung in pJOE2702 durchgeführt:

Ca. 50 ng Produkt aus PCR 1.1 und ca. 50 ng Produkt aus PCR1.2 mit
0,25 µM MAT350 (5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' SEQ ID NO
58), 0,25 µM MAT353 (5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' SEQ ID NO
25 61), 100 µM dNTP, 10 µL Pfu-Polymerase-Puffer (10x), 2,5 U Pfu-Polymerase
(Zugabe bei 85 °C, "hot start") und H₂O ad 100 µL.

Temperaturprofil:

1. 95°C 10 min, 2. 85 °C 5 min, 3. 59 °C 30 s, 4. 72 °C 2 min, 5. 95 °C 30
30 s, 6. 72°C 10 min
Zyklen: (1-2.) 1x, (3-5.) 22x, (6.) 1x

Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Ligation in Vektor pPCR-Script-Amp, eine Klonierung in *Escherichia coli* XL1-Blue, Sequenzierung der Insertion, Spaltung mit NdeI und BamHI sowie eine Ligation in pJOE2702. Das erhaltene Plasmid
5 wurde als pBT4 bezeichnet.

Charakterisierung und Nachweis der Enzymaktivität von CarB (Phytoendesaturase)

Das von *carB* abgeleitete Genprodukt wurde als CarB bezeichnet. CarB
10 weist auf Grundlage der Peptidsequenzanalyse folgende Eigenschaften auf:

	Länge:	582 Aminoacylreste
	Molekulare Masse:	66470
	Isoelektrische Punkt:	6,7
15	Katalytische Aktivität:	Phytoendesaturase
	Edukt:	Phytoen
	Produkt:	Lycopin
	EC-Nummer:	EC 1.14.99-

Der Nachweis der Enzymaktivität erfolgte *in vivo*. Wenn das Plasmid
20 (pCAR-AE) in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE). Dieser Stamm synthetisiert Phytoen. Wenn zusätzlich das Plasmid pBT4 in *Escherichia coli* XL1-Blue übertragen wird, entsteht der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4). Da ausgehend von *carB* eine enzymatisch aktive
25 Phytoendesaturase gebildet wird, produziert dieser Stamm Lycopin.

Die Plasmide pCAR-AE und pBT4 wurden daher in *Escherichia coli* übertragen. Nach Wachstum in Flüssigkultur wurden die Carotinoide aus den Zellen extrahiert und charakterisiert (vgl. oben).

Durch HPLC Analyse wurde nachgewiesen, daß der Stamm *Escherichia coli* XL1-Blue (pCAR-AE) Phytoen und der Stamm *Escherichia coli* XL1-Blue (pCAR-AE)(pBT4) Lycopin produziert. CarB weist folglich die Enzymaktivität einer Phytoendesaturase auf.

5

Herstellung gentechnisch veränderter Stämme von *Blakeslea trispora* zur Herstellung von Phytoen

Nachfolgend werden beispielhaft die Herstellung von gentechnisch veränderten Organismen zur Herstellung von Phytoen beschrieben.

10

Vector pBinAHyg Δ carB zur Erzeugung von carB⁻ -Mutanten von *Blakeslea trispora*

Für die Deletion von carB in *Blakeslea trispora* wurde der Vektor pBinAHyg Δ carB (SEQ. ID. NO:62, Fig. 22) konstruiert. Der Vorläufer von pBinAHyg Δ carB ist pBinAHyg (SEQ. ID. NO:3, Fig. 2). pBinAHyg wurde folgendermaßen konstruiert:

15

20

25

Aus dem Plasmid pANsCos1 (SEQ. ID. NO:4, Fig. 1, Osiewacz, 1994, Curr. Genet. 26:87-90) wurde die gpdA-hph Kasette als BglII/HindIII Fragment isoliert und in das BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet und enthält das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors und des trpC Terminators aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind.

Die Amplifikation der codierenden Sequenz von carB mit den Primern MAT350 und MAT353 mittels PCR wurde mit den folgenden Parametern durchgeführt:

30

50 ng pBT4 mit 0,25 μ M MAT350 (5'-ACTTTATTGGATCCTTAAAT-GCGAATATCGTTGCTGC-3'; SEQ ID NO 58), 0,25 μ M MAT353 (5'-

CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3'; SEQ ID NO 61), 100 µM dNTP, 10 µL Pfu-Polymerase-Puffer, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µL H₂O

Temperaturprofil:

- 5 1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 30s, 4. 72°C 2 min, 5. 95 °C 30s, 6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

- 10 Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Spaltung mit HindIII, eine weitere Reinigung des 364-bp-HindIII-Fragments-carB, gefolgt von einer Spaltung von pBinAHyg mit HindIII, eine Ligation von 364-bp-HindIII-Fragments-carB in pBinAHyg, eine Transformation des Vektors in *Escherichia coli* und eine Isolierung des Konstruktes und Bezeichnung als pBinAHygΔcarB wie oben beschrieben.
- 15 Alternativ erfolgte eine partielle Spaltung mit HindIII und die Klonierung eines größeren HindIII-Fragmentes aus carB in pBinAHyg zur Herstellung von pBinAHygΔcarB.

Erzeugung von carB⁻ -Mutanten von *Blakeslea trispora*

- 20 Zunächst wurde das Plasmid pBinAHygΔcarB in den Agrobakterienstamm LBA 4404 übertragen, z. B. durch Elektroporation (vgl. oben). Anschließend wurde das Plasmid von *Agrobacterium tumefaciens* LBA 4404 in *Blakeslea trispora* ATCC 14272 und in *Blakeslea trispora* ATCC 14271 übertragen (vgl. oben). Der erfolgreiche Nachweis des
- 25 Gentransfers in *Blakeslea trispora* erfolgte über Polymerase-Kettenreaktion nach folgendem Protokoll:

- Ca. 0,5 µg DNA aus *Blakeslea trispora* ATCC 14272 carB⁻ bzw. ATCC 14271 carB⁻ wurden mit 0,25 µM Primer hph forward (5'-CGATGTAGGAGGGCGTGGATA-3'; SEQ ID NO 5), 0,25 µM Primer hph
- 30 reverse (5'-GCTTCTGCGGGCGATTTGTGT-3'; SEQ ID NO 6), 100 µM

dNTP, 10 µL Herculase-Polymerase-Puffer, 2,5 U Herculase-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µl H₂O umgesetzt.

Temperaturprofil:

1. 95°C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1
5 min, 6. 72°C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

Als Negativkontrolle wurde eine Amplifikation des Kanamycinresistenzgens aus *Agrobacterium* versucht. Dazu wurden
10 folgende PCR-Bedingungen verwendet:

Ca. 0,5 µg DNA aus *Blakeslea trispora* ATCC 14272 carB⁻ bzw. ATCC 14271 carB⁻ wurden mit 0,25 µM Primer nptIII forward (5'-TGAGAATATCACCGGAATTG-3'; SEQ ID NO 7), 0,25 µM Primer nptIII reverse (AGCTCGACATACTGTTCTTCC-3'; SEQ ID NO 8), 100 µM
15 dNTP, 10 µL Herculase-Polymerase-Puffer, 2,5 U Herculase-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 µL H₂O umgesetzt.
Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min-

20 Zyklen: (1-2.) 1x, (3-5.) 30x, (6.) 1x

Produktion von Carotinoiden und Carotinoidvorstufen mit *Blakeslea trispora*

25 Zur Produktion der Carotinoide Zeaxanthin, Canthaxanthin, Astaxanthin und Phytoen wurden die entsprechenden gentechnisch veränderten *Blakeslea trispora* (+) und (-) Stämme fermentiert, das produzierte Carotinoid mittels HPLC Analyse nachgewiesen und isoliert.

Das Flüssigmedium zur Produktion von Carotinoiden enthielt pro Liter: 19 g Maismehl, 44 g Sojamehl, 0,55 g KH_2PO_4 , 0,002 g Thiaminhydrochlorid, 10 % Sonnenblumenöl. Der pH wurde mit KOH auf 7,5 eingestellt.

- 5 Zur Herstellung der Carotinoiden wurden Schüttelkolben mit Sporensuspensionen von (+) und (-) Stämmen der GVO von *Blakeslea trispora* beimpft. Die Schüttelkolben wurden bei 26 °C mit 250 rpm für 7 Tage inkubiert. Alternativ wurde zu Mischungen der Stämme nach 4 Tagen Trisporsäuren zugegeben und weitere 3 Tage inkubiert. Die
10 Endkonzentration der Trisporsäuren betrug 300 - 400 µg/ml.

Extraktion und Analytik

Extraktion:

1. Entnahme von 10 ml Kultursuspension
- 15 2. Zentrifugation, 10 min, 5.000 x g
3. Verwerfen des Überstandes
4. Resuspendierung des Pellets in 1 ml Tetrahydrofuran (THF) durch Vortexen
5. Zentrifugation, 5 min, 5.000 x g
- 20 6. Abnahme der THF-Phase
7. Wiederholung der Schritte 4.-6. (2 x)
8. Vereinigung der THF-Phasen
9. Zentrifugation der vereinigten THF-Phasen 5 min bei 20.000 x g, um Reste der wäßrigen Phase abzutrennen

25

Analytik

Messung von Phytoen mittels HPLC

- Säule: ZORBAX Eclipse XDB-C8, 5 µm, 150*4,6 mm
- Temperatur: 40 °C
- 30 Flußrate: 0,5 ml/min
- Injektionsvolumen: 10 µl

Detektion: UV 220 nm

Stoppzeit: 12 min

Nachlaufzeit: 0 min

Maximaldruck: 350 bar

5 Eluent A: 50 mM NaH₂PO₄, pH 2,5 mit Perchlorsäure

Eluent B: Acetonitril

Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
	0	50	50	0,5
10	12	50	50	0,5

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 50 - 1000 mg/l eingewogen und in THF gelöst. Als Standard wurde Phytoen verwendet, welches unter den gegebenen Bedingungen eine Retentionszeit von 7,7 min. aufweist.

Messung von Lycopin, β-Carotin, Echinenon, Canthaxanthin, Cryptoxanthin, Zeaxanthin und Astaxanthin mittels HPLC

Säule: Nucleosil 100-7 C18, 250*4,0 mm (Macherey & Nagel)

Temperatur: 25 °C

Flußrate: 1,3 ml/min

Injektionsvolumen: 10 µl

25 Detektion: 450 nm

Stoppzeit: 15min

Nachlaufzeit: 2 min

Maximaldruck: 250 bar

Eluent A: 10% Aceton, 90% H₂O

30 Eluent B: Aceton

Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
	0	30	70	1,3
	10	5	95	1,3
	12	5	95	1,3
5	13	30	70	1,3

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden
 10 jeweils 10 mg eingewogen und in 100 ml THF gelöst. Als Standard wurden folgende Carotinoide mit folgenden Retentionszeiten eingesetzt β-Carotin (12,5 min), Lycopin (11,7 min), Echinenon (10,9 min), Cryptoxanthin (10,5 min), Canthaxanthin (8,7 min), Zeaxanthin (7,6 min) und Astaxanthin (6,4 min) [s. Fig 23].

15

Produktion von Zeaxanthin mit gentechnisch veränderten Stämmen von *Blakeslea trispora*

Nachfolgend wird beispielhaft die Herstellung von Zeaxanthin mit gentechnisch veränderten Organismen (GVO) von *Blakeslea trispora*
 20 beschrieben.

Durch Agrobakterium-vermittelte Transformation wurde der Vektor pBinAHygBTpTEF1-HPcrtZ in *Blakeslea trispora* übertragen (s.o.). Ein Hygromycin-resistenter Klon wurde isoliert und auf eine Kartoffel-Glucose-Agarplatte (Merck KGaA, Darmstadt) übertragen.

25 Nach drei Tagen Inkubation bei 26°C wurde ausgehend von dieser Platte eine Sporensuspension hergestellt. Ein 250-ml-Erlenmeyerkolben ohne Schikanen mit 50 ml Growth-Medium (Maismehl 47 g/l, Sojamehl 23 g/l, KH₂PO₄ 0,5 g/l, Thiamin-HCl 2.0 mg/l, pH mit NaOH vor der Sterilisation auf 6,2–6,7 eingestellt) wurde mit 1x10⁵ Sporen beimpft.
 30 Diese Vorkultur inkubierte 48 Stunden bei 26 °C und 250 upm. Für die Hauptkultur wurde ein 250-ml-Erlenmeyerkolben ohne Schikane

enthaltend 40 ml Produktionsmedium mit 4 ml der Vorkultur beimpft und 8 Tage bei 26 °C und 150 upm inkubiert. Das Produktionsmedium enthielt Glucose 50 g/l, Casein Acid Hydrolysate 2 g/l, Hefeextrakt 1 g/l, L-Asparagin 2 g/l, KH_2PO_4 1,5 g/l, $\text{MgSO}_4 \times 7 \text{ H}_2\text{O}$ 0,5 g/l, Thiamin-HCl 5
5 mg/l, Span20 10 g/l, Tween 80 1 g/l, Linolsäure 20 g/l, Maisquellwasser 80 g/l. Nach 72 Stunden erfolgte die Zugabe von Kerosin in einer Endkonzentration von 40 g/l Kerosin.

Nach der Ernte der Kulturen werden die verbliebenen ungefähr 35 ml Kultur mit Wasser auf 40 ml aufgefüllt. Anschließend werden die Zellen im
10 Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gaulin, 3 x bei 1500 bar aufgeschlossen.

Die Suspension mit den aufgeschlossenen Zellen wurde mit 35 ml THF versetzt und 60 min bei RT im Dunkeln bei 250 upm geschüttelt. Danach wurden 2 g NaCl zugegeben und das Gemisch nochmals geschüttelt. Der
15 Extraktionsansatz wurde dann 10 min bei 5000 x g zentrifugiert. Die gefärbte THF-Phase wurde abgenommen, die Zellmasse war vollständig entfärbt.

Die THF-Phase wurde am Rotationsverdampfer bei 30 mbar und 30 °C auf 1 ml eingeeengt und danach nochmals in 1 ml THF aufgenommen.
20 Nach Zentrifugation 5 min bei 20 000 x g wurde ein Aliquot der oberen Phase entnommen und durch HPLC analysiert (Fig. 24, Fig. 23).

Patentansprüche

1. Verfahren zur Herstellung eines gentechnisch veränderten Organismus der Gattung *Blakeslea* umfassend
 - 5 (i) Transformation mindestens einer der Zellen,
 - (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder mehreren genetischen Merkmalen alle gleichartig verändert sind und diese gentechnische Veränderung zur Ausprägung bringen, und
 - 10 (iii) Selektion und Anzucht der gentechnisch veränderten Zelle oder Zellen.
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** es sich um Zellen von Pilzen der Art *Blakeslea trispora* handelt.
3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** in
 - 15 der Transformation (i) ein Vector oder freie Nukleinsäuren verwendet werden.
4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector in das Genom mindestens einer der Zellen integriert wird.
- 20 5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector einen Promotor und/oder einen Terminator enthält.
6. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 5, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector
 - 25 enthaltend den *gpd*, *pcarB*, *pcarRA* und/oder *ptef1* Promotor und/oder den *trpC* Terminator eingesetzt wird.

7. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 6, **dadurch gekennzeichnet, dass** in der Transformation (i) ein Vector enthaltend ein Resistenzgen eingesetzt wird.
- 5 8. Verfahren nach Anspruch 7, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector ein Hygromycin-Resistenzgen (hph), insbesondere aus *E. coli* enthält.
9. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der gpd Promotor die Sequenz SEQ ID NO: 1
10 aufweist.
10. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch gekennzeichnet, dass** der trpC Terminator die Sequenz SEQ ID NO: 2 aufweist.
11. Verfahren nach einem der vorhergehenden Ansprüche 5 - 8, **dadurch**
15 **gekennzeichnet, dass** der tef1 Promotor die Sequenz SEQ ID NO: 35 aufweist.
12. Verfahren nach einem der Ansprüche 6 bis 11, **dadurch gekennzeichnet, dass** der gpd Promotor und der trpC Terminator aus *Aspergillus nidulans* stammen.
- 20 13. Verfahren nach einem Ansprüche 3 bis 12, **dadurch gekennzeichnet, dass** der Vector die SEQ ID NO: 3 umfasst.
14. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Transformation (i) mittels Agrobakterien, Konjugation, Chemikalien, Elektroporation, Beschuss mit DNA-beladenen Partikeln, Protoplasten oder Mikroinjektion durchgeführt
25 wird.

15. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** in der Homokaryontisierung (ii) ein mutagenes Agens eingesetzt wird.
16. Verfahren nach Anspruch 15, **dadurch gekennzeichnet, dass** als
5 mutagenes Agens N-Methyl-N'-nitro-nitrosoguanidin (MNNG), UV-Strahlung oder Röntgenstrahlung eingesetzt wird.
17. Verfahren nach einem der vorhergehenden Ansprüche, **dadurch gekennzeichnet, dass** die Selektion durch Markierung und/oder Auswahl der einkernigen Zellen erfolgt.
- 10 18. Verfahren nach einem der vorhergehenden Ansprüche 1 - 17, **dadurch gekennzeichnet, dass** in der Selektion 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) oder 5-Fluororotat (FOA) und Uracil und Hygromycin eingesetzt werden.
- 15 19. Verfahren nach einem der Ansprüche 3 bis 18, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder deren Vorstufen enthält.
- 20 20. Verfahren nach einem der Ansprüche 3 bis 19, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinen oder Xanthophyllen enthält.
- 25 21. Verfahren nach einem der Ansprüche 3 bis 20, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β -Carotin, α -Carotin, Lutein, Bixin, Phytofluen oder Phytoen enthält.

22. Verfahren nach einem der Ansprüche 3 bis 21, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector derart gestaltet ist, dass die im Vector enthaltene genetische Information in das Genom von *Blakeslea trispora* eingeführt wird
- 5 23. Verfahren nach einem der Ansprüche 3 bis 22, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector genetische Informationen enthält, die nach Expression eine Ketolase- und/oder Hydroxylase-Aktivität entfalten.
- 10 24. Verfahren nach Anspruch 23, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector die SEQ ID NO: 70 oder 71 oder 76 und/oder 72 umfasst.
25. Verfahren nach Anspruch 23 oder 24, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector eine Sequenz aus der Gruppe bestehend aus den SEQ ID NO: 37 – 51 aufweist.
- 15 26. Verfahren nach einem der Ansprüche 3 bis 21, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector derart gestaltet ist, dass die im Vector enthaltene genetische Information in der Zelle ausgeschaltet wird.
- 20 27. Verfahren nach einem der Ansprüche 3 bis 21 oder 25, **dadurch gekennzeichnet, dass** durch die Transformation (i) das Gen der Phytoendesaturase ausgeschaltet wird.
28. Verfahren nach Anspruch 27, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector die **SEQ ID NO: 69** umfasst.
- 25 29. Verfahren nach Anspruch 27 oder 28, **dadurch gekennzeichnet, dass** der in der Transformation (i) eingesetzte Vector die Sequenz SEQ ID NO: 62 aufweist.

30. Verfahren nach einem der Ansprüche 3 bis 21, **dadurch gekennzeichnet, dass** durch die Transformation das Gen der Lycopincyclase ausgeschaltet wird.
- 5 31. Genetisch veränderte mehrkernige Zellen der zur Gattung *Blakeslea* gehörenden Pilze, insbesondere *Blakeslea trispora* erhältlich nach einem der vorhergehenden Ansprüche.
32. Verwendung der Zellen nach Anspruch 30 oder eines aus ihnen gebildeten Mycels zur Herstellung von Carotinoiden oder deren Vorstufen.
- 10 33. Verwendung nach Anspruch 30 oder 31 zur Herstellung von Carotinen oder Xanthophyllen.
34. Verwendung nach einem der Ansprüche 30 bis 32 zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-
15 Hydroxyechinenon, Lycopin, β -Carotin, α -Carotin, Lutein, Bixin, Phytofluon oder Phytoen.
35. Promotor mit der Sequenz SEQ ID NO: 1 oder 35 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 29.
36. Terminator mit der Sequenz SEQ ID NO: 2 zur Verwendung in dem
20 Verfahren nach einem der Ansprüche 1 - 29.
37. Vector umfassend die SEQ ID NO: 3 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 29.
38. Vector nach Anspruch 36 zur Verwendung in dem Verfahren nach einem der Ansprüche 1 - 29 umfassend die SEQ ID NO: 69 und/oder
25 die SEQ ID NO: 70 oder 71 und/oder 72 oder 76.

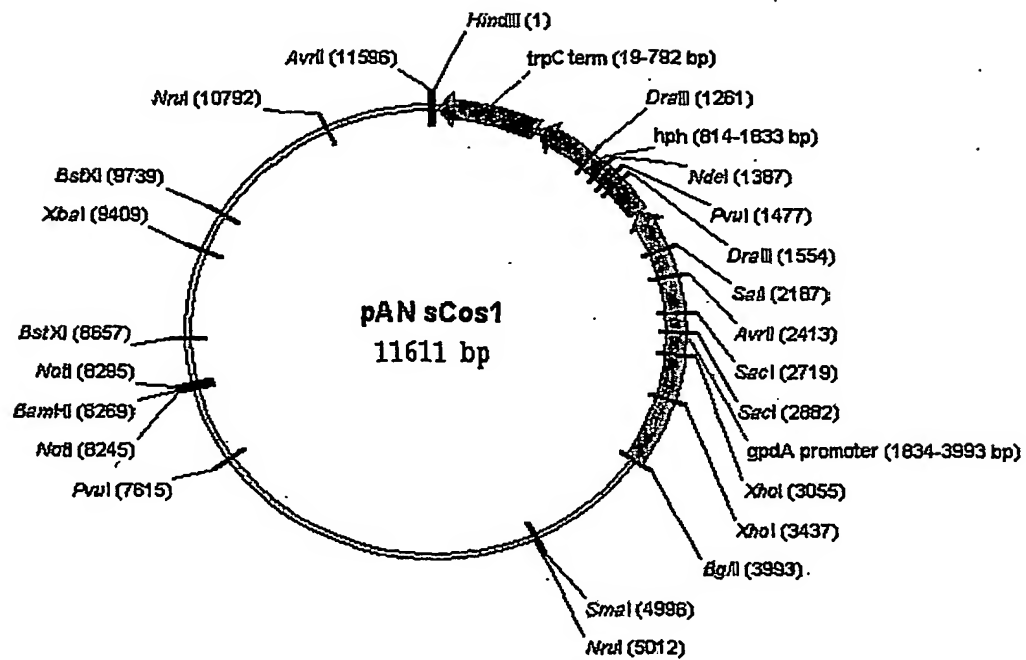
Fig. 1: Vektor pANsCos1

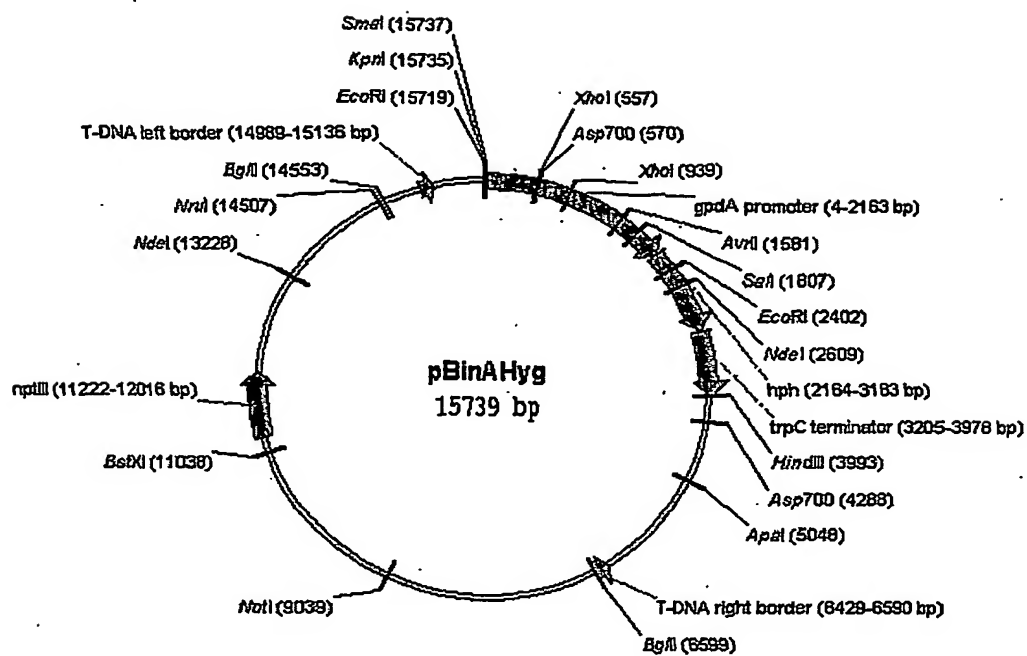
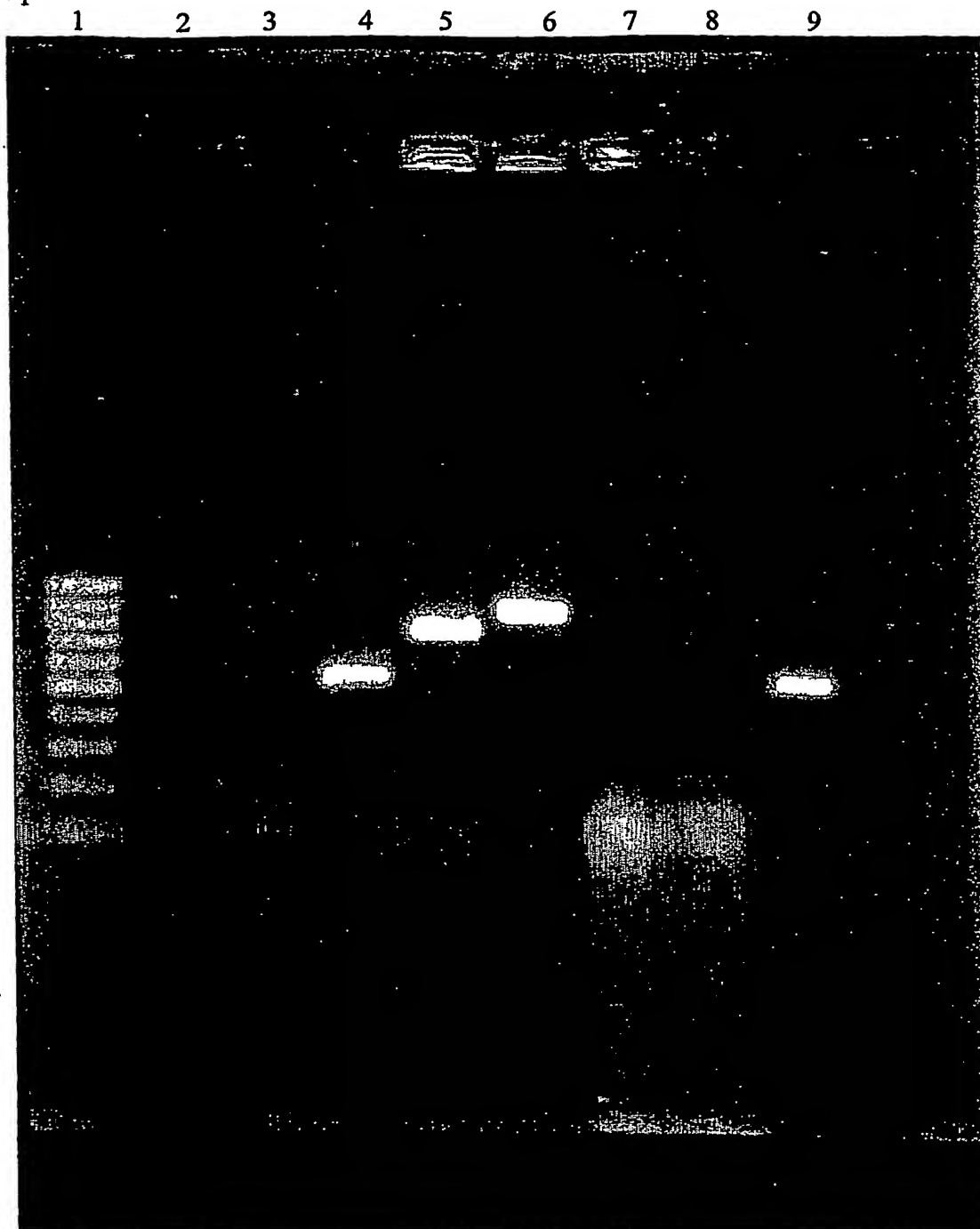
Fig. 2: Vektor pBinAHyg

Fig. 3: Gels des Ergebnis einer PCR
Spur:



BEST AVAILABLE COPY

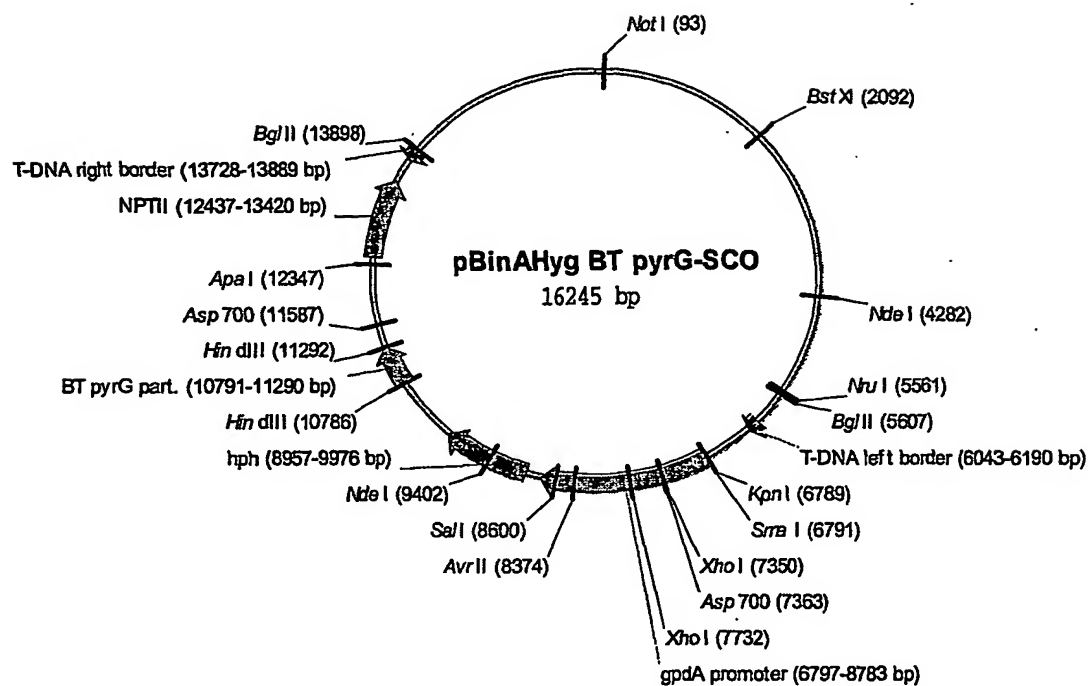
Fig. 4: Plasmid pBinAHygBTpyrG-SCO

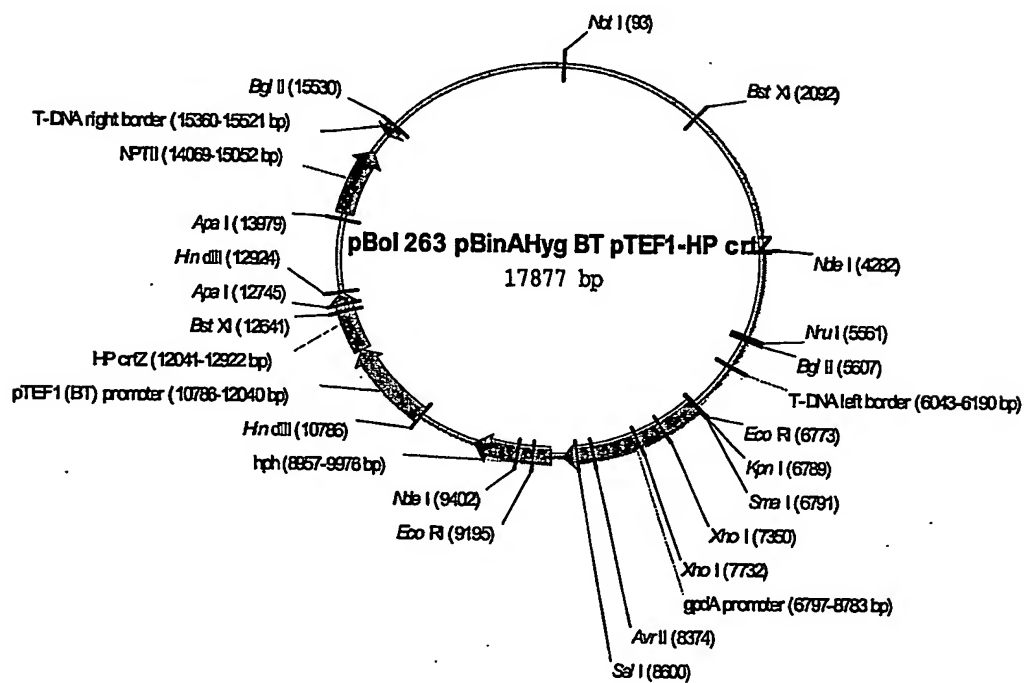
Fig. 5: Plasmid pBinAHygBTpTEF1-HPcrtZ

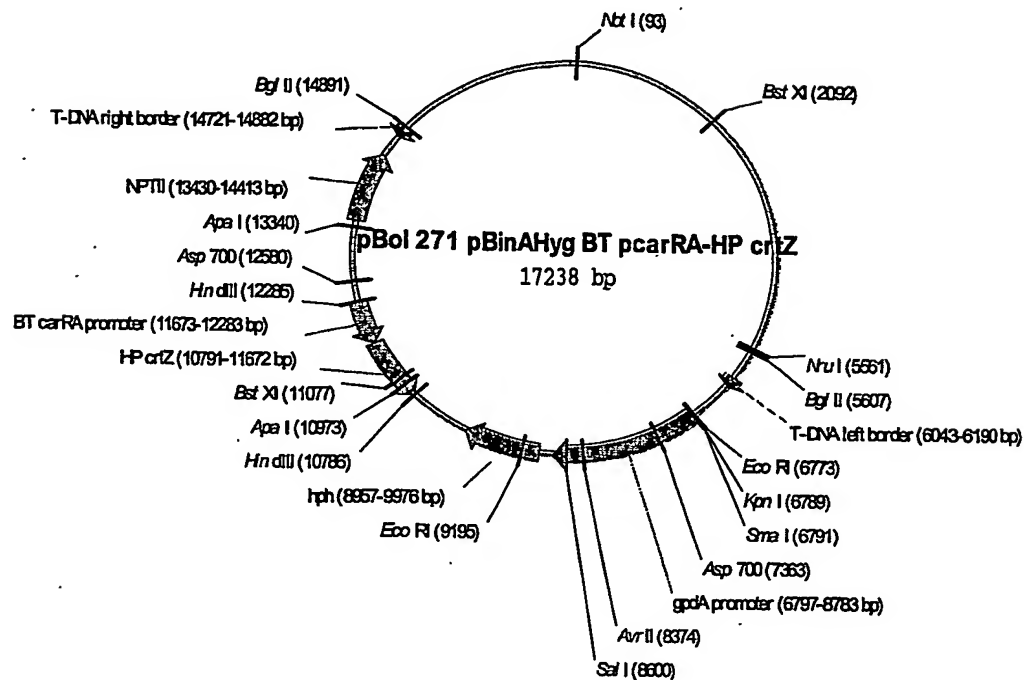
Fig. 6: Plasmid pBinAHyg-BTpcarRA-HPcrtZ

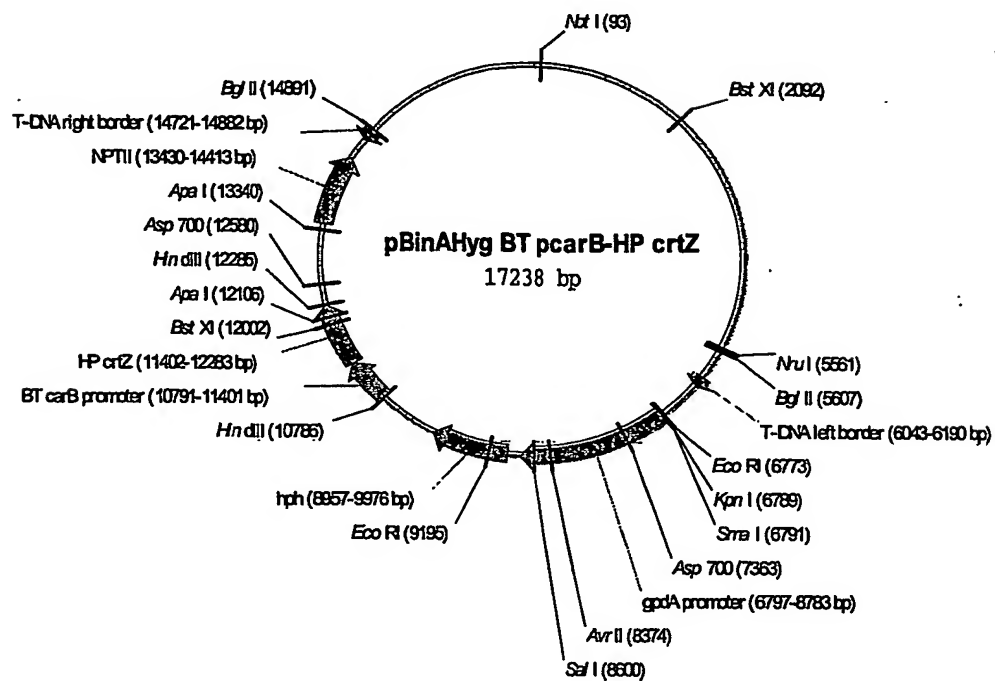
Fig. 7: Plasmid pBinAHygBTpcarB-HPcrtZ

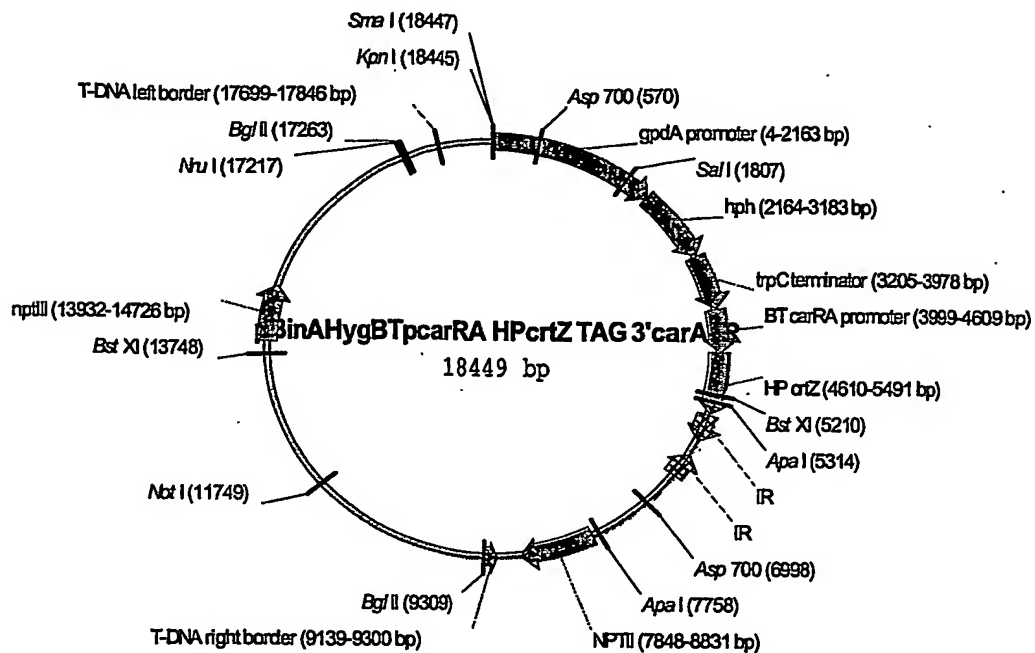
Fig. 8: Plasmid p-carRA-HPcrtZ-TAG-3'carA-IR

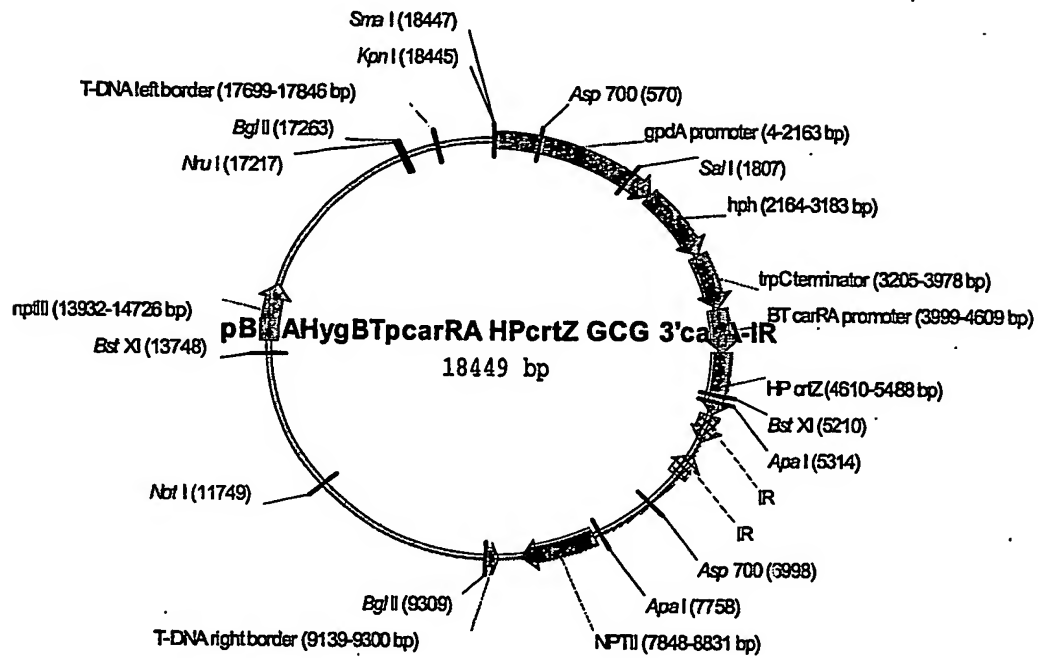
Fig. 9: Plasmid p-carRA-HPcrtZ-GCG-3'carA-IR

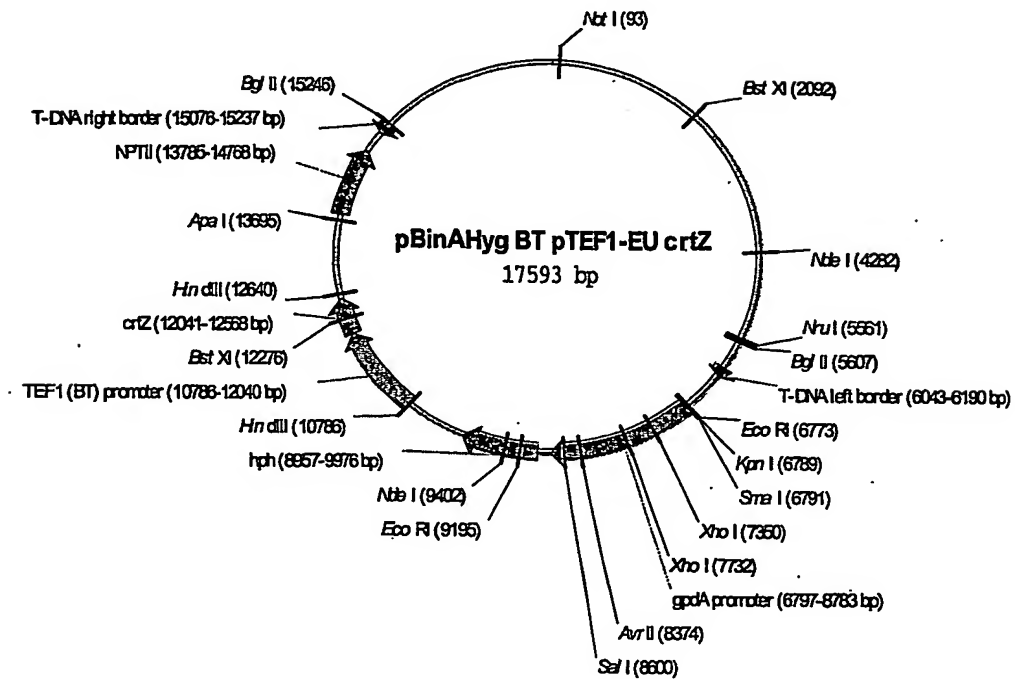
Fig. 10: Plasmid pBinAHygBTpTEF1-EUcrtZ

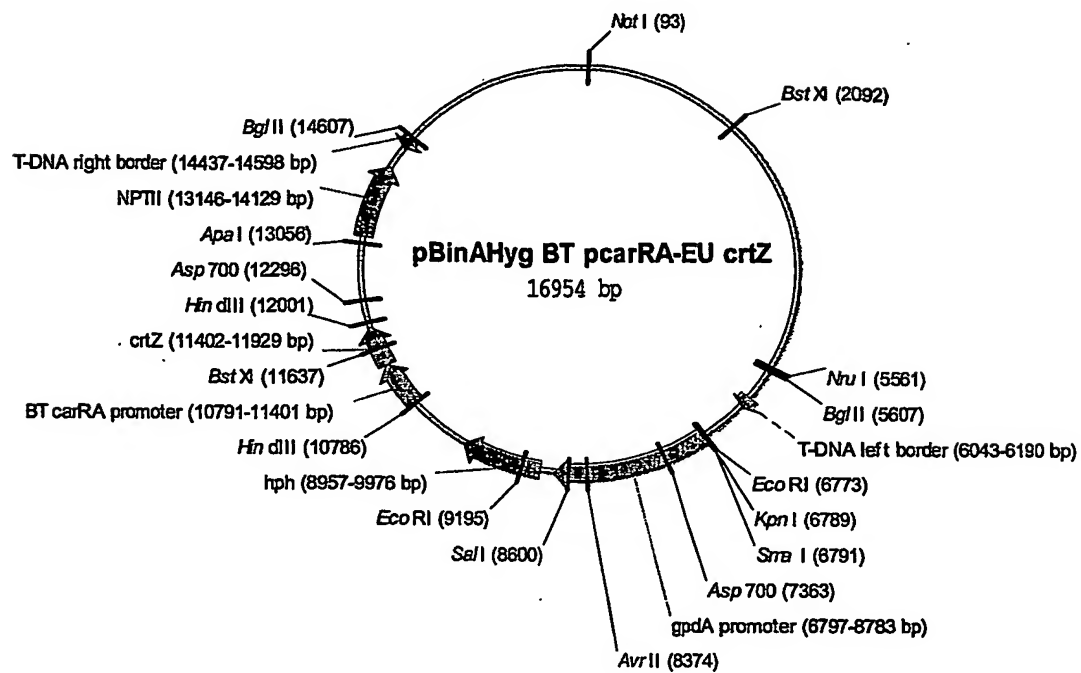
Fig. 11: Plasmid pBinAHygBTpcarRA-EUcrtZ

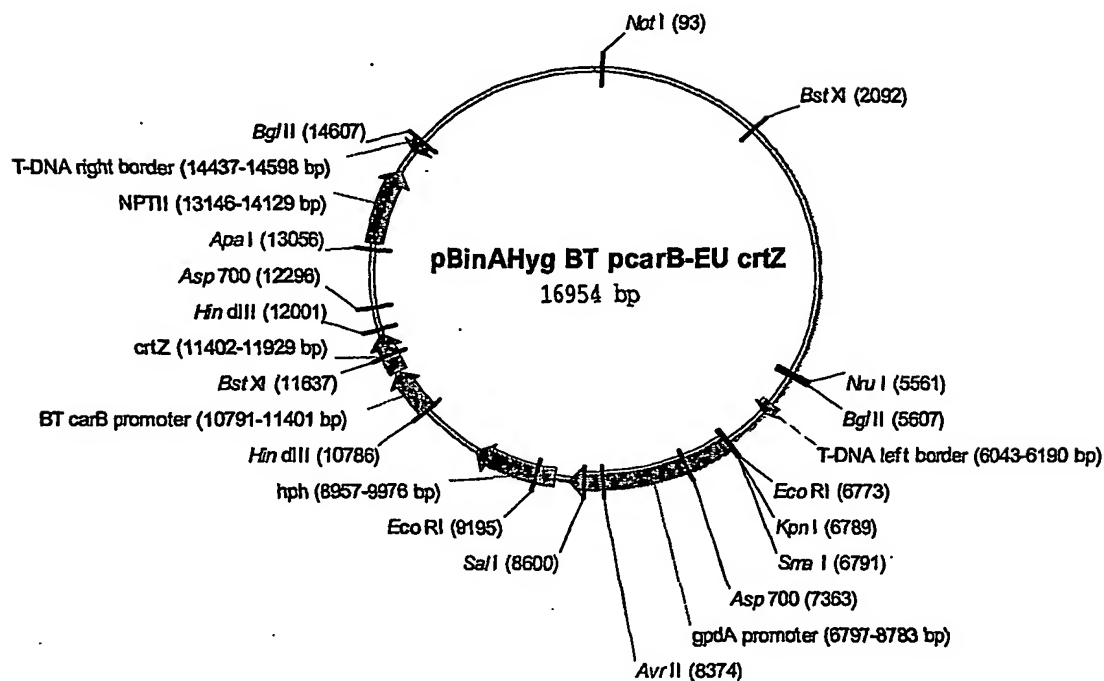
Fig. 12: Plasmid pBinAHygBTpcarB-EUcrtZ

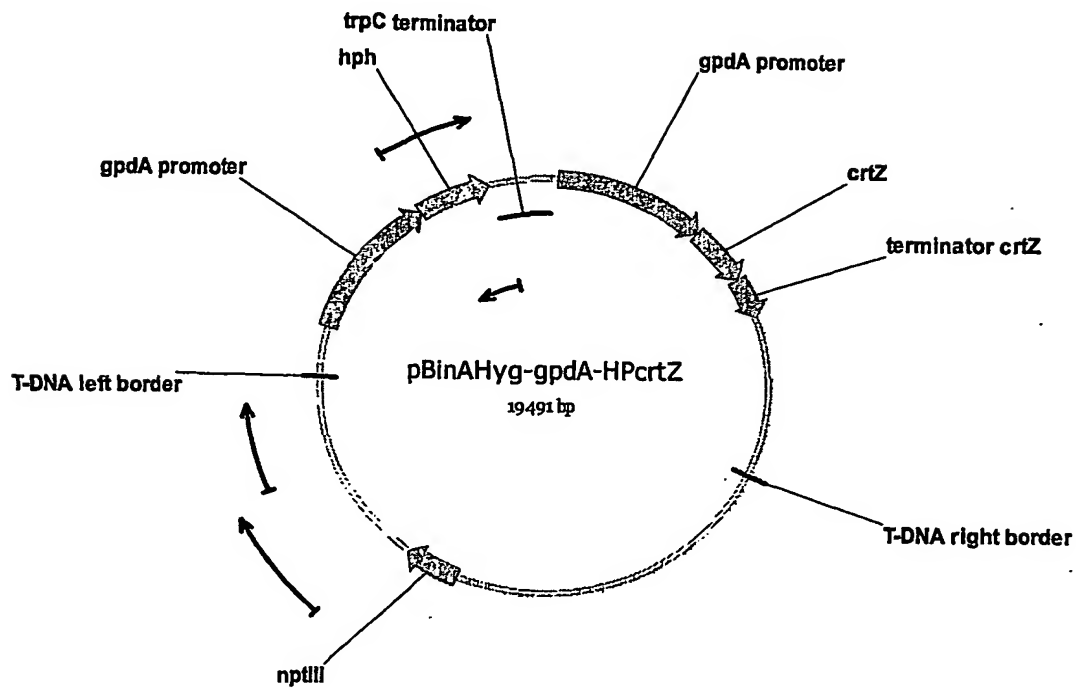
Fig. 13: Plasmid p-BinAHyg-gpdA-HPcrtZ

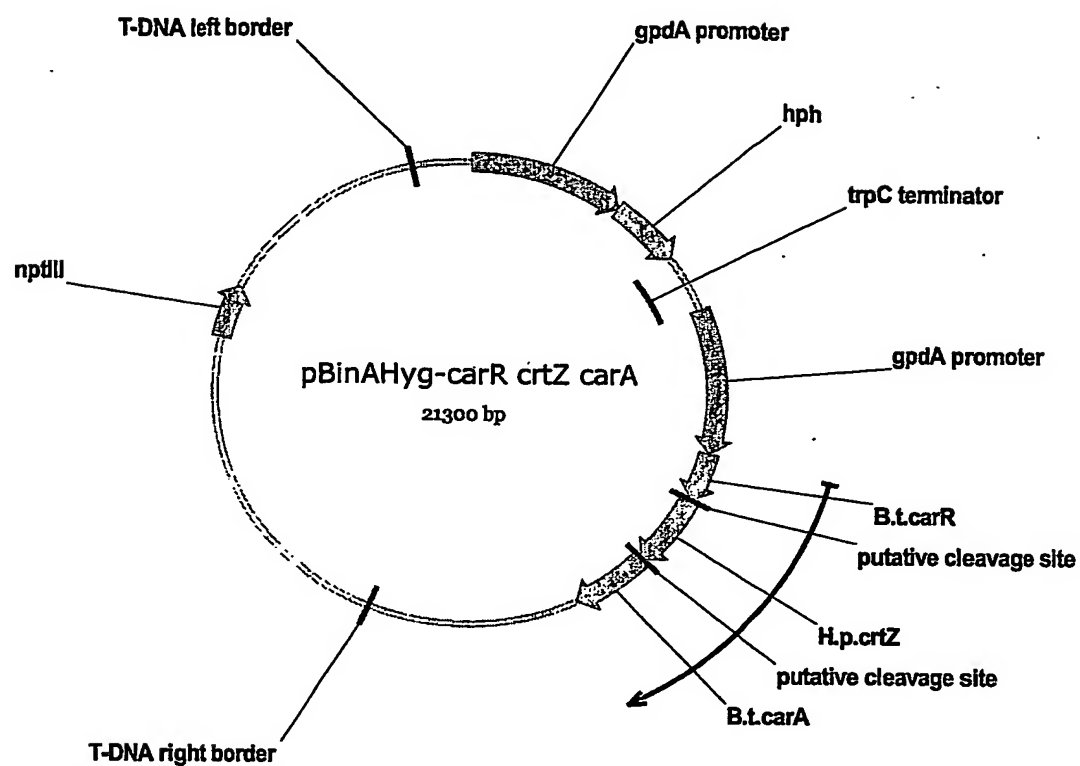
Fig. 14: Plasmid pBinAHyg-carRcrtZcarA

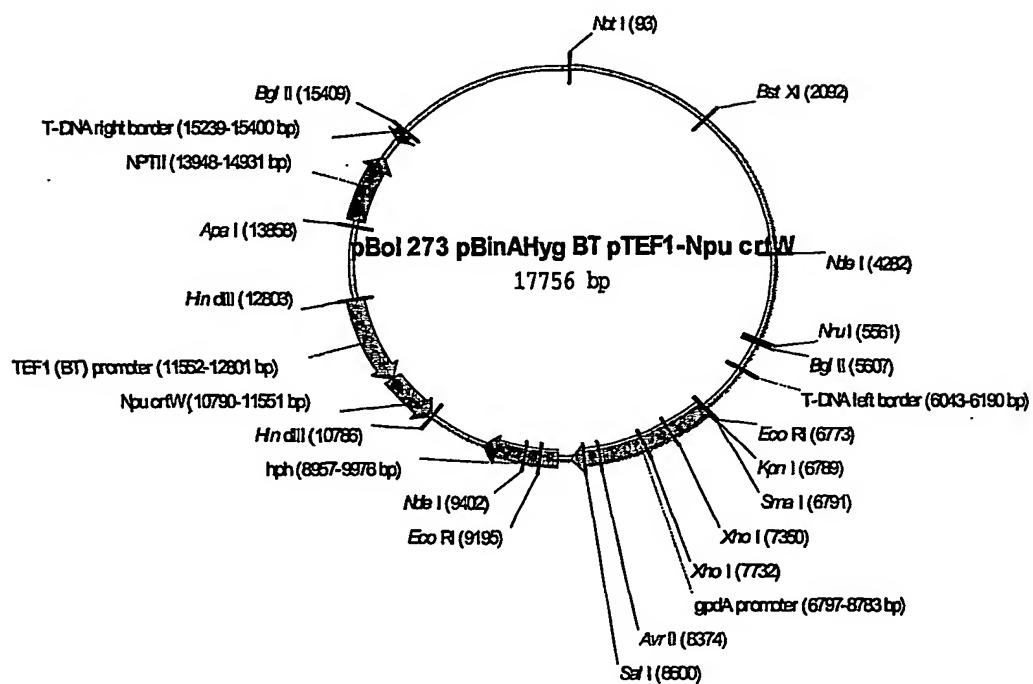
Fig. 15: Plasmid pBinAHyg-BTpTEF1-NPcrtW

Fig. 16: Plasmid pBinAHyg_BTpcarRA_NPcrtW

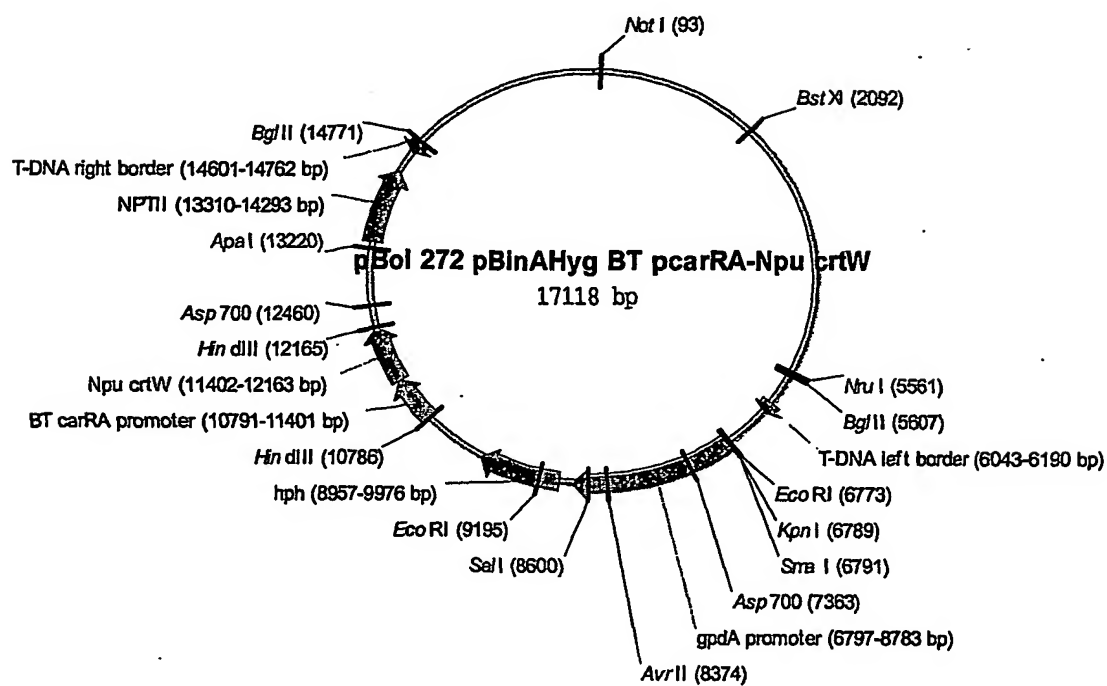


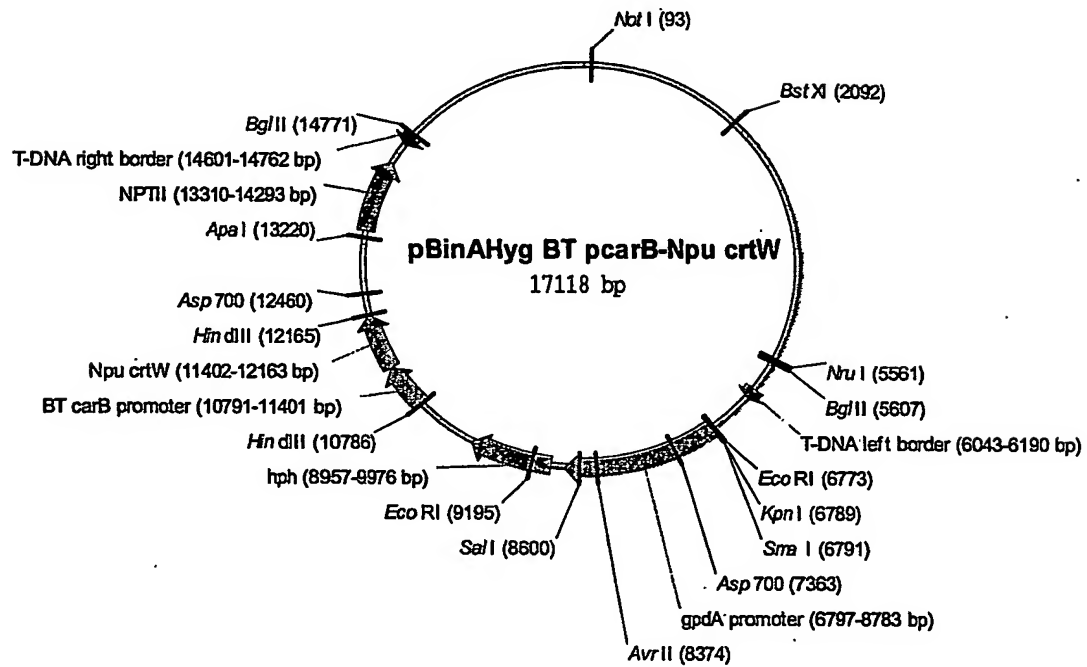
Fig. 17: Plasmid pBinAHyg-BTpcarB-NPcrtW

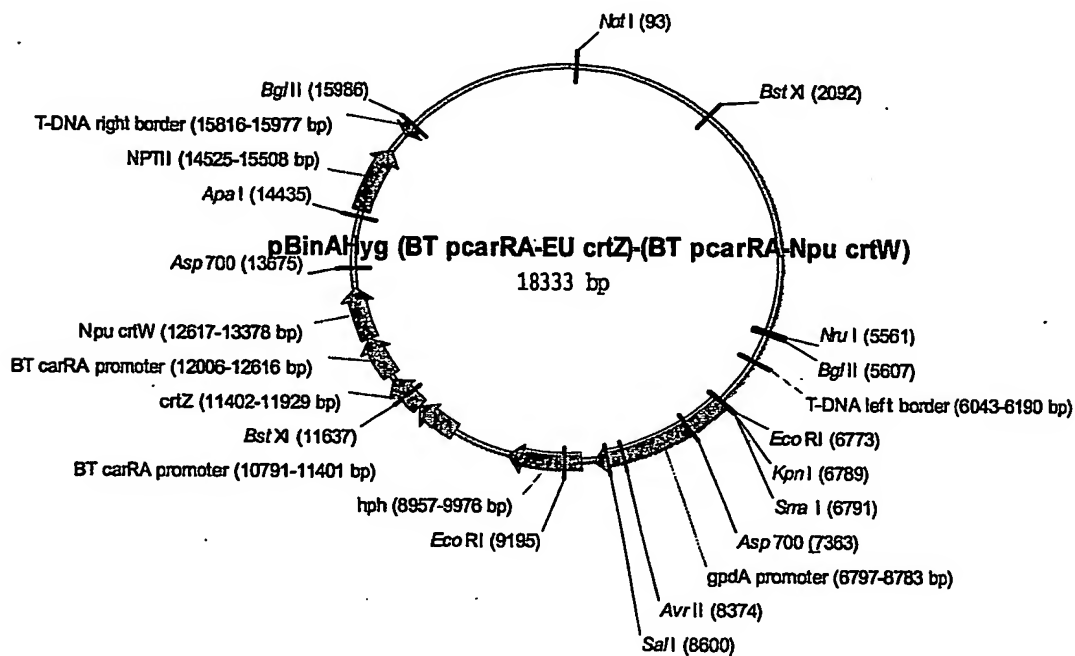
Fig. 19: Plasmid pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW

Fig. 20: carB

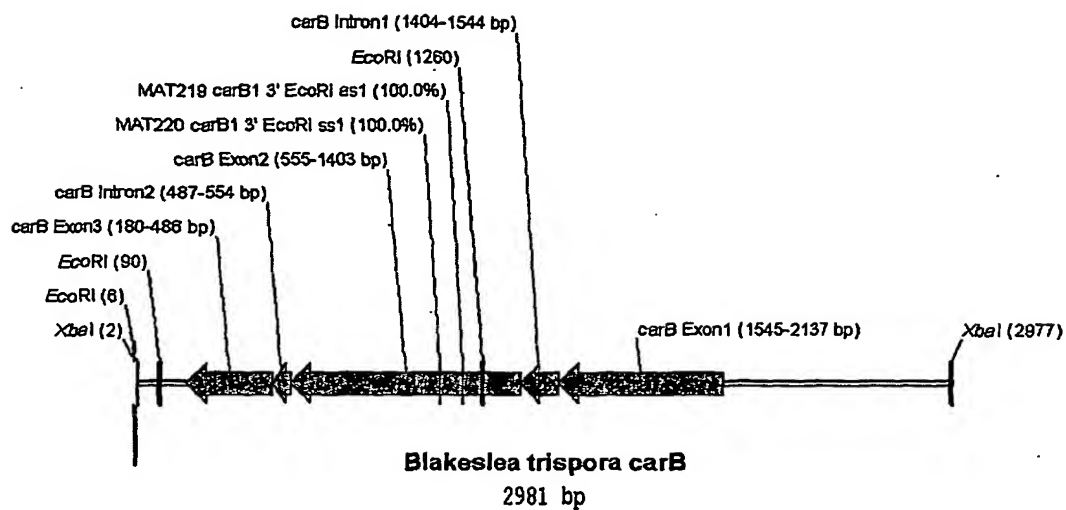


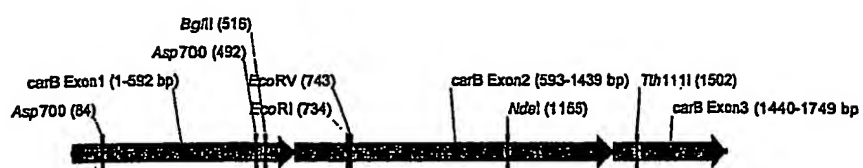
Fig. 21: CDS von carB

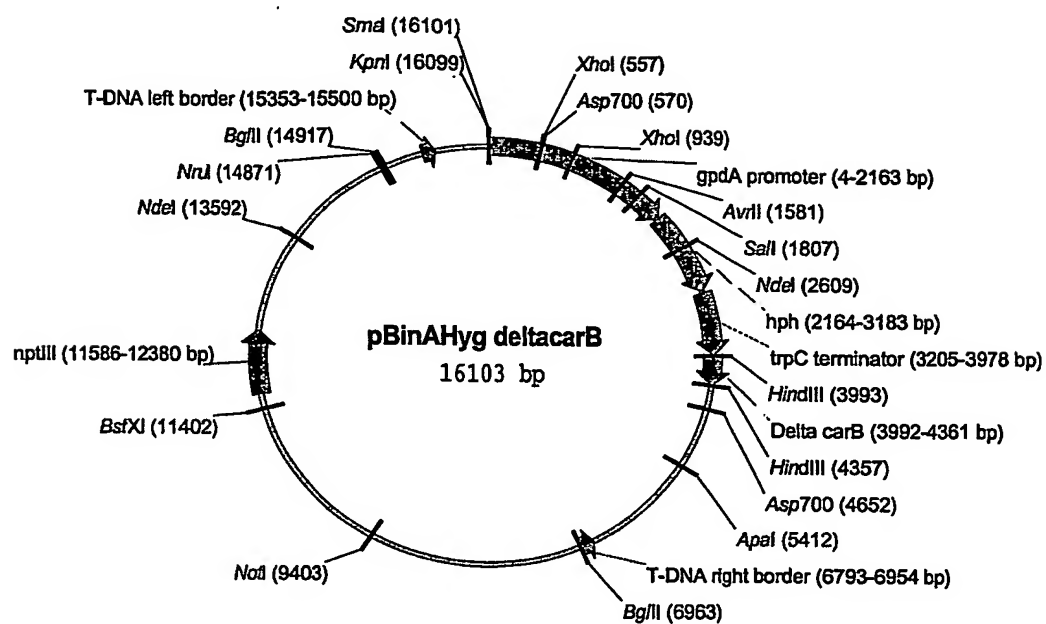
Fig. 22: Vektor pBinAHyg Δ carB

Fig. 23: HPLC Standard

Sequence File C:\HPCHEM\2\SEQUENCE\CAROTIN.S
Method C:\HPCHEM\2\METHODS\CAROTIN.M
Last changed 11/12/2003
Carotinoide Nucleoside 11 5100-7µ C18 250*4,0mm
DAD1 A, Slp=450.20 Ret=off (131103.1) PROBE018.D

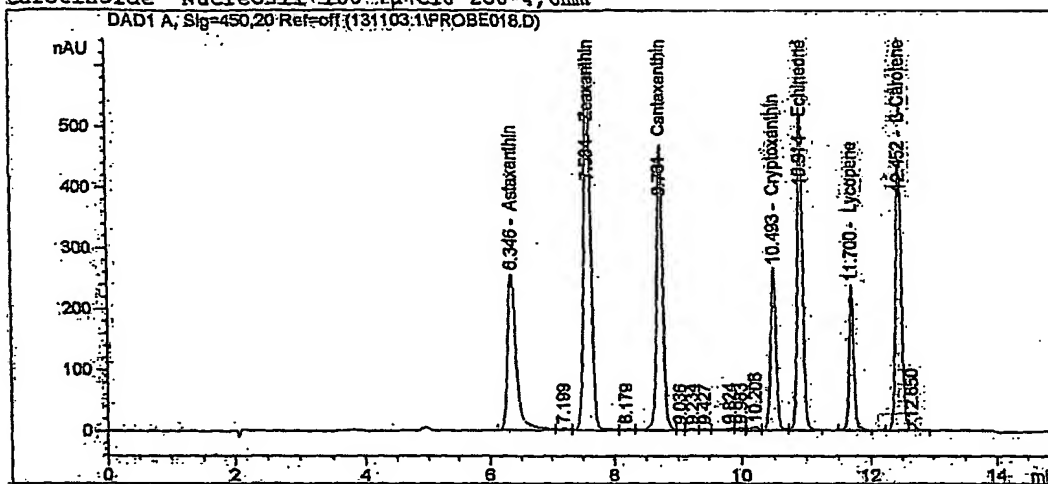
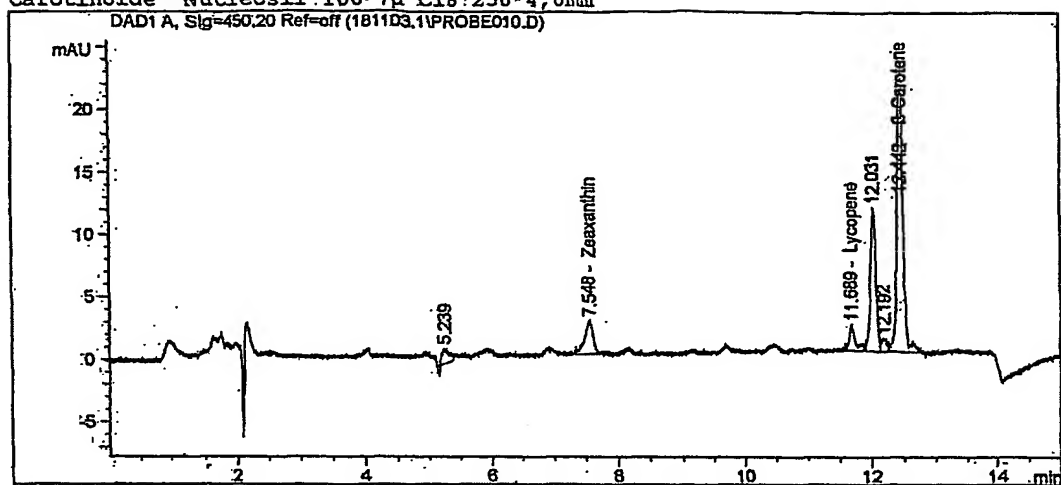


Fig. 24: HPLC

Sequence File : C:\HPCHEM\2\SEQUENCE\CAROTIN.S
Method : C:\HPCHEM\2\METHODS\CAROTIN.M
Last changed : 11/18/2003
Carotinoide Nucleosil 100-7µ C18: 250*4, 0mm
DAD1 A, Slg=450.20 Ref=off (181103,1VPROBE010.D)



SEQUENCE LISTING

<110> BASF AG

<120> Verfahren zur gentechnischen Veränderung von Organismen der
Gattung Blakeslea, entsprechende Organismen und deren Verwendung

<130> ?

<160> 80

<170> PatentIn version 3.2

<210> 1

<211> 2160

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 1

ctttcgacac tgaaatacgt cgagcctgct ccgcttggaa gggcgagga gcctcgtcct	60
gtcacaacta ccaacatgga gtacgataag ggccagttcc gccagtcac taagagccag	120
ttcatgggag ttggcatgat ggccgtcatg catctgtact tcaagtacac caacgctctt	180
ctgatccagt cgatcatccg ctgaaggcgc tttcgaatct ggttaagatc cacgtcttcg	240
ggaagccagc gactggtgac ctccagcgtc cctttaagga tgccaacagc tttctcagcc	300
agggccagcc caagaccgac aaggcctccc tccagaacgc cgagaagaac tggaggggtg	360
gtgtcaagga ggagtaagct ccttattgaa gtcggaggac ggagcgggtg caagaggata	420
ttcttcgact ctgtattata gataagatga tgaggaattg gaggtagcat agcttcattt	480
ggatttgctt tccaggctga gactctagct tggagcatag agggtccttt ggctttcaat	540
attctcaagt atctcgagtt tgaacttatt ccctgtgaac cttttattca ccaatgagca	600

ttggaatgaa catgaatctg aggactgcaa tcgccatgag gttttcgaaa tacatccgga 660

tgtcgaaggc ttggggcacc tgcgttggtt gaatttagaa cgtggcacta ttgatcatcc 720

gatagctctg caaagggcgt tgcacaatgc aagtcaaacg ttgctagcag ttccaggtgg 780

aatgttatga tgagcattgt attaaatcag gagatatagc atgatctcta gttagctcac 840

cacaaaagtc agacggcgta accaaaagtc acacaacaca agctgtaagg atttcggcac 900

ggctacggaa gacggagaag ccaccttcag tggactcgag taccatttaa ttctatttgt 960

gtttgatcga gacctaatc agccctaca acgaccatca aagtcgtata gctaccagtg 1020

aggaagtgga ctcaaatac cttcagcaac atctcctgga taaactttta gcctaaacta 1080

tacagaataa gataggtgga gagcttatac cgagctccca aatctgtcca gatcatgggt 1140

gaccggtgcc tggatcttcc tatagaatca tccttattcg ttgacctagc tgattctgga 1200

gtgaccaga gggcatgac ttgagcctaa aatccgccgc ctccaccatt tgtagaaaaa 1260

tgtgacgaac tcgtgagctc tgtacagtga ccggtgactc tttctggcat gcggagagac 1320

ggacggacgc agagagaagg gctgagtaat aagccactgg ccagacagct ctggcggctc 1380

tgaggtgcag tggatgatta ttaatccggg accggccgcc cctccgcccc gaagtggaaa 1440

ggctggtgtg cccctcgttg accaagaatc tattgcatca tcggagaata tggagcttca 1500

tcgaatcacc ggcagtaagc gaaggagaat gtgaagccag ggggtgtatag ccgtcggcga 1560

aatagcatgc cattaaccta ggtacagaag tccaattgct tccgatctgg taaaagattc 1620

acgagatagt accttctccg aagtaggtag agcgagtacc cggcgcgtaa gctccctaata 1680

tggcccatcc ggcattctgta gggcgtccaa atatcgtgcc tctcctgctt tgcccggtgt 1740

atgaaaccgg aaaggccgct caggagctgg ccagcggcgc agaccgggaa cacaagctgg 1800

cagtcgaccc atccggtgct ctgcactcga cctgctgagg tccctcagtc cctggtaggc 1860

agctttgccc cgtctgtccg cccggtgtgt cggcgggggtt gacaaggctg ttgcgtcagt 1920
ccaacatttg ttgccatatt ttccgtgtct cccaccagc tgctcttttc ttttctcttt 1980
cttttcccat cttcagtata ttcattcttc catccaagaa cctttatttc ccctaagtaa 2040
gtactttgct acatccatac tccatccttc ccatccctta ttcccttgaa cctttcagtt 2100
cgagctttcc cacttcacg cagcttgact aacagctacc ccgcttgagc agacatcacc 2160

<210> 2

<211> 774

<212> DNA

<213> Artificial

<220>

<223> Terminator

<220>

<221> misc_feature

<222> (267)..(267)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (475)..(475)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (566)..(566)

<223> n is a, c, g, or t

<400> 2

cgatccactt aacgttactg aaatcatcaa acagcttgac gaatctggat ataagatcgt 60
tggtgtcgat gtcagctccg gagttgagac aaatggtgtt caggatctcg ataagatacg 120
ttcatttgtc caagcagcaa agagtgcctt ctagtgattt aatagctcca tgtcaacaag 180

aataaaacgc gttttcgggt ttacctcttc cagatacagc tcatctgcaa tgcattaatg 240
cattgactgc aacctagtaa cgccttncag gtcgccgga agagaagaat agcttagcag 300
agctattttc attttcggga gacgagatca agcagatcaa cggtcgtcaa gagacctacg 360
agactgagga atccgctctt ggctccacgc gactatatat ttgtctctaa ttgtactttg 420
acatgctcct cttctttact ctgatagctt gactatgaaa attccgtcac cagcncctgg 480
gttcgcaaag ataattgcat gtttcttcct tgaactctca agcctacagg acacacattc 540
atcgtaggta taaacctcga aatcanttcc tactaagatg gtatacaata gtaaccatgc 600
atggttgctt agtgaatgct ccgtaacacc caatacgccg gccgaaactt ttttacaact 660
ctcctatgag tcgtttaccc agaatgcaca ggtacacttg tttagaggta atccttcttt 720
ctagctagaa gtccctcgtgt actgtgtaag cgcccactcc acatctccac tcga 774

<210> 3

<211> 15739

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 3

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt	60
cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct	240
tccgggaagcc agcgactggg gacctccagc gtccctttaa ggctgccaac agcttttctca	300
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggaattg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggc cgttgcacia tgcaagtcaa acgttgctag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg	900
cacggctacg gaagacggag aagccacctt cagtggactc gagtaccait taattctatt	960
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa	1080
ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg	1140

gttgaccggg gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgacct agagggcat gacttgagcc taaaatccgc cgctccacc attttagaa 1260

aatgtgacg aactcgtgag ctctgtacag tgaccgggta ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gcccctccgc cccgaagtgg 1440

aaaggctggg gtgcccctcg ttgaccaaga atctattgca tcctcgagaa atatggagct 1500

tcctcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtcgaatt gcttccgac tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggg gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctg ccgcccggg tgctggcggg gttgacaagg tcgttgcgtc 1920

agtccaacat ttgttgccat attttctgc tctcccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccaattca tcgcagcttg actaacagct acccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcggt gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgatcc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcatc tcccgccgtg cacaggggtgt cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgtgtt ctgcagccgg tcgcgagggc catggatgcg 2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
ggccaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
tggcaaatg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga tttcggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccgg agcttgcagg atcgccgcgg 2940
ctccgggcgt atatgctccg cattggctct gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgccccg agaagcgcgg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctcgctccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttgggtg cgatgtcagc tccggagtgt agacaaatgg 3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ctttctagtg 3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaacctg gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggctcc acgcgactat 3600
atatttgtct ctaattgtac tttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattcgc tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780
gatggtatac aatagtaacc atgcatgggt gcctagttaa tgctccgtaa cacccaatac 3840
gccggccgaa actttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900
cttggtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccactcgacc tgcaggcatg caagcttggc gtaatcatgg tcatagctgt 4020
ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa 4080
agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac 4140
tgcccgcctt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 4200
cggggagagg cggtttgcgt attgggcca agacaaaagg ggcacattca accgattgag 4260
ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca ccgtcaccga 4320
cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4380
ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta ggcacagaat 4440
caagtttgcc tttagcgta gactgtagcg cgttttcatc ggcattttcg gtcatagccc 4500
ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg 4560
aaccgcctcc ctgagagccg ccacctcag aaccgccacc ctgagagcca ccacctcag 4620
agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggagge ccgatctagt 4680
aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct 4740
atcgcgattt aaatgtataa ttgcgggact ctaatcataa aaacctatct cataaataac 4800
gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4860
atcatcgcaa gaccggcaac aggattcaat ctttaagaaac tttattgcca aatgtttgaa 4920

cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4980
atatcgcggt gcatctcggt cttgcctggg cagtgcgcgc cgacgccgtt gatgtggacg 5040
ccggggccga tcatattgtc gctcaggatc gtggcgttgt gcttgtcggc cgttgctgtc 5100
gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc gaagaactcc 5160
agcatgagat cccgcgcgtg gaggatcatc cagccggcgt cccggaaaac gattccgaag 5220
cccaaccttt catagaaggc ggcgggtggaa tcgaaatctc gtgatggcag gttgggcgtc 5280
gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca agaaggcgat 5340
agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5400
cccattcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg tcctgatagc 5460
ggtcgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca 5520
tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc 5580
gcgccttgag cctggcgaac agttcgggtg gcgcgagccc ctgatgctct tcgtccagat 5640
catcctgacg gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5700
cttggtggtc gaatgggcag gtagccgat caagcgtatg cagccgccgc attgcatcag 5760
ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca 5820
cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc 5880
aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctgctcctgc agttcattca 5940
gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga 6000
acacggcggc atcagagcag ccgattgtct gttgtgcccc gtcatagccg aatagcctct 6060
ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc 6120
cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 6180

gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg 6240

accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6300

actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg 6360

taaaacggct tgtcccgct catcgccggg ggtcataacg tgactccctt aattctccgc 6420

tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg acaggatata 6480

ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat taaaagggc 6540

gtgaaaagg t tatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga 6600

tctggcgccg gccagcgaga cgagcaagat tggccgccc cgaacgat ccgacagcgc 6660

gccagcaca ggtgcgagg caaattgcac caacgcatac agcgccagca gaatgccata 6720

gtggcggtg acgtcgttcg agtgaaccag atcgcgagg agggccggca gcaccggcat 6780

aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat 6840

gttgggttcc acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggtt 6900

ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6960

acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggtcggc 7020

gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca gccggcgctt 7080

tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg 7140

gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt tctgatcggg 7200

aatgcccgca gcttcaggca ggcgctgtc gcctaccgag atggcgcgcg catccatgcc 7260

ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg ctctctctgc 7320

gaggcggtt tttcgccgg ggacgccgtc aatgcgctga tgacaatcag ctacttcact 7380

gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc 7440

gttgaacagg ctccgctctc gccgctgttg cgggccgcga tagacgcctt cgacgaagcc 7500
ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg 7560
aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg 7620
ccggagcgca acccactcac tacagcagag ccatgtagac aacatccccct ccccctttcc 7680
accgcgtcag acgcccgtag cagcccgcta cgggcttttt catgccctgc cctagcgtcc 7740
aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttctctcg 7800
tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 7860
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 7920
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 7980
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 8040
gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 8100
ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgcttttcc 8160
gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc atcctttttc 8220
gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttggtg tatccaacgg 8280
cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc cttcttctact 8340
gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccgggt 8400
accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8460
ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag 8520
gcggcgggcg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa 8580
atcacggggc tcgtggacta tgagcacgtc cgcgagctgg cccgcacaa tggcgacctg 8640
ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac ggcgcggttc 8700

ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8760
aaggtcatga tgggcgtggt ccgcccagg gcagagccat gactttttta gccgctaaaa 8820
cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8880
cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga 8940
cgctcaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc cctgcaaacg 9000
cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt tgtggatacc 9060
tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact tgaggggccc 9120
actcaccggc cgccggcgtt acagatgagg ggcaggctcg atttcggccg gcgacgtgga 9180
gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc ccacagatga 9240
tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9300
acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgaggggccc 9360
acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc aagggtttcc 9420
gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9480
accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg aaggggggtg 9540
cccccccttc tcgaaccctc ccggcccgtt aacgcgggcc tccatcccc ccaggggctg 9600
cgcccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9660
gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc cggaagcatt 9720
gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9780
ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga cttcatggcg 9840
gggccggcaa tttttacctt gggcattott ggcatagtgg tcgcgggtgc cgtgctcgtg 9900
ttcgggggtg cgataaacc agcgaaccat ttgaggatgat aggtaagatt ataccgaggt 9960

atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat ttaaaaagct 10020
accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat attgacaata 10080
ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 10140
caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg 10200
catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat 10260
tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc gatgactttg 10320
tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc 10380
tgctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct gattacgtgc 10440
agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca tatcaccag 10500
tcaaagggtg acagcaggct cataagacgc cccagcgctg ccatagtgcg ttcaccgaat 10560
acgtgogcaa caaccgtctt ccggagactg tcatacgctg aaaacagcca gcgctggcgc 10620
gatttagccc cgacatagcc ccactgttcg tccatttccg cgcagacgat gacgtcactg 10680
cccggctgta tgcgcgaggt taccgactgc ggctgagtt ttttaagtga cgtaaaatcg 10740
tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca ttcattggcca 10800
tatcaatgat tttctggtgc gtaccggggt gagaagcggg gtaagtgaac tgcagttgcc 10860
atgttttacg gcagtgcgag cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10920
gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10980
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 11040
gtttcaaaat cggtccgctc gatactatgt tatacgccaa ctttgaaaac aactttgaaa 11100
aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 11160
tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 11220

aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11280

aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag aaaatgaaaa 11340

cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11400

aaaggacatg atgctatggc tggaaggaaa gctgcctggt ccaaagggtcc tgcactttga 11460

acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc tttgctcgga 11520

agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcacag 11580

gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11640

agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11700

agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11760

cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct ttgtgaaaga 11820

tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggatatga 11880

cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt atgtcgagct 11940

atTTTTTgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 12000

ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 12060

ccgacttctt ccgcatcaag tgTTTTTggt ctCaggccga ggcccacggc aagtatttgg 12120

gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac gagaaggacg 12180

gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12240

caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc ggggcaatcc 12300

cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12360

acgcggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc atgCGTgcgc 12420

cccgcgaaac cttccagtcc gtcggctcga tggTccagca agctacggcc aagatcgagc 12480

gcgacagcgt gcaactggct cccctgccc tgcccgcgcc atcgccgcc gtggagcgtt 12540
cgcgctgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc gacacgcgag 12600
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12660
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12720
ccttgttcga tattgcgcgcg tggccggaca cgatgcgagc gatgccaaac gacacggccc 12780
gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg caaaacaagg 12840
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcgag ctgcgggccc 12900
acgatgacga actgggtgtg cagcaggtgt tggagtacgc gaagcgcacc cctatcggcg 12960
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg atcaatggcc 13020
ggatttacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg atgggcttca 13080
cgtcgacccg cgttgggcac ctggaatcgg tgctgctgct gcaccgcttc cgcgtcctgg 13140
accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc gtcgtgctgt 13200
ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13260
cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgctc aagctggaaa 13320
ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc gagcaggtcg 13380
gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg gtcaatgatg 13440
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg ggttcagcag 13500
ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact tgcttcgctc 13560
agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag gattaaaatt 13620
gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13680
cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg ttacagagca 13740

cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg tggcattcgg 13800

cgctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg acggcccca 13860

ggacgctcac aaggcgcac tgctccggcg tttcgtggag ccggaacagc gaggccgagg 13920

ggtcgccggt atgctgctgc gggcggtgcc ggcggtttta ttgctcgtga tgatcgtccg 13980

acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac ttaatatctc 14040

gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg tcgcgcgac 14100

ggtaggcgt gtgcagccgc tgatggctgt gttcatctct gccgctctgc taggtagccc 14160

gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg cgctgttggt 14220

gttgacacca aacgcagcgc tagatcctgt cggcgctgca gcgggcctgg cgggggcggt 14280

ttccatggcg ttcggaaccg tgctgaccgc caagtggcaa cctcccgctc ctctgctcac 14340

ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag ctttagtggt 14400

tgatccgcca atcccgatgc ctacaggaac caatgttctc ggctggcgt ggctcggcct 14460

gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14520

agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc cggggatgca 14580

tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtag caatggatag 14640

gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc ttcctcagcg 14700

gctttatcca gcgatttcct attatgtcgg catagtcttc aagatcgaca gcctgtcacg 14760

gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgagg agatgatatt 14820

tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14880

atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14940

gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 15000

ctgtatcgag tgggtgatttt gtgccgagct gccggtcggg gagctggttg ctggctggtg 15060
gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac acattgcgga 15120
cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa cagctgattg 15180
cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggt ttgccccagc 15240
aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat aaatcaaaag 15300
aatagcccga gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga 15360
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 15420
aaccatcacc caaatcaagt tttttggggt cgagggtgccg taaagcacta aatcggaacc 15480
ctaaaggag ccccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 15540
aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15600
tcgggtcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 15660
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtga 15720
ttcgagctcg gtacccggg 15739

<210> 4

<211> 11611

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (227)..(227)

<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (318)..(318)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (526)..(526)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (8946)..(8946)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10028)..(10028)
<223> n is a, c, g, or t

<400> 4
agcttgcacg cctgcaggtc gaggaggat gtggaggagg cgcttacaca gtacacgagg 60
acttctagct agaaagaagg attacctcta aacaagtgtg cctgtgcatt ctgggtaaag 120
gactcatagg agagttgtaa aaaagtttcg gccggcgatg tgggtgttac ggagcattca 180
ctaggcaacc atgcattggtt actattgtat accatcttag taggaantga tttcgagggtt 240
tataacctacg atgaatgtgt gtcctgtagg cttgagagtt caaggaagaa acatgcaatt 300
atctttgcga acccaggngc tgggtgacgga attttcatag tcaagctatc agagtaaaga 360
agaggagcat gtcaaagtac aattagagac aaatatatag tcgcgtggag ccaagagcgg 420
attcctcagt ctcttaggtc tcttgacgac cgttgatctg cttgatctcg tctccgaaa 480
atgaaaatag ctctgctaag ctattcttct cttcgccgga gcctgnaagg cgttactagg 540
ttgcagtcaa tgcattaatg cattgcagat gagctgtatc tggaagaggt aaaccgaaa 600
acgcgtttta ttcttggtga catggagcta ttaaactcact agaaggcact ctttgctgct 660

tggacaaatg aacgtatctt atcgagatcc tgaacaccat ttgtctcaac tccggagctg	720
acatcgacac caacgatctt atatccagat tcgtcaagct gtttgatgat ttcagtaacg	780
ttaagtggat cgatcccgcg gtcggcatct actctattcc tttgccctcg gacgagtgt	840
ggggcgctcg tttccactat cggcgagtac ttctacacag ccatcgggtcc agacggccgc	900
gcttctgcgg gcgatttgtg tacgcccgcg agtcccggct ccggatcgga cgattgcgtc	960
gcatcgaccc tgcgccaag ctgcatcatc gaaattgccg tcaaccaagc tctgatagag	1020
ttggtcaaga ccaatgcgga gcatatacgc ccggagccgc ggcgatcctg caagctccgg	1080
atgcctccgc tcgaagtagc gcgtctgtg ctccatacaa gccaaaccacg gcctccagaa	1140
gaagatgttg gcgacctcgt attgggaatc cccgaacatc gcctcgctcc agtcaatgac	1200
cgctgttatg cggccattgt ccgtcaggac attgttggag ccgaaatccg cgtgcacgag	1260
gtgccggact tcggggcagt cctcggccca aagcatcagc tcatcgagag cctgcgcgac	1320
ggacgcactg acggtgtcgt ccatcacagt ttgccagtga tacacatggg gatcagcaat	1380
cgcgcatatg aaatcacgcc atgtagtgta ttgaccgatt ccttgcggtc cgaatgggcc	1440
gaaccgcgtc gtctggctaa gatcggccgc agcgatcgca tccatggcct ccgcgaccgg	1500
ctgcagaaca gcgggcagtt cggtttcagg caggctcttc aacgtgacac cctgtgcacg	1560
gcgggagatg caataggtca ggctctcgct gaattcccca atgtcaagca cttccggaat	1620
cgggagcgcg gccgatgcaa agtgccgata aacataacga tctttgtaga aaccatcggc	1680
gcagctattt acccgcagga catatccacg ccctcctaca tcgaagctga aagcacgaga	1740
ttcttcgccc tccgagagct gcatcaggtc ggagacgctg tcgaactttt cgatcagaaa	1800
cttctcgaca gacgtcgcgg tgagttcagg catggtgatg tctgctcaag cggggtagct	1860
gttagtcaag ctgcgatgaa gtgggaaagc tcgaactgaa aggttcaaag gaataaggga	1920

tgggaaggat ggagtatgga tgtagcaaag tacttactta ggggaaataa aggttcttgg 1980

atgggaagat gaatatactg aagatgggaa aagaaagaga aaagaaaaga gcagctggtg 2040

gggagagcag gaaaatatgg caacaaatgt tggactgacg caacgacctt gtcaacccccg 2100

ccgacacacc gggcggacag acggggcaaa gctgcctacc agggactgag ggacctcagc 2160

aggtcgagtg cagagcaccg gatgggtcga ctgccagctt gtgttcccgg tctgcgccgc 2220

tggccagctc ctgagcggcc tttccggttt catacaccgg gcaaagcagg agaggcacga 2280

tatttgagcgc ccctacagat gccggatggg ccaattaggg agcttacgcg ccgggtactc 2340

gctctaccta cttcggagaa ggtactatct cgtgaatctt ttaccagatc ggaagcaatt 2400

ggacttctgt acctaggtta atggcatgct atttcgccga cggctataca cccctggctt 2460

cacattctcc ttcgcttact gccggtgatt cgatgaagct ccatattctc cgatgatgca 2520

atagattctt ggtcaacgag gggcacacca gcctttccac ttcggggcgg aggggcggcc 2580

ggtcccggat taataatcat ccaactgcacc tcagagccgc cagagctgtc tggccagtgg 2640

cttattactc agcccttctc tctgcgtccg tccgtctctc cgcatgccag aaagagtcac 2700

cggtcactgt acagagctca cgagttcgtc acatttttct acaaattggtg gaggcggcgg 2760

attttaggct caagtcatga ccctctgggt cactccagaa tcagctaggt caacgaataa 2820

ggatgattct ataggaagat ccaggcaccg gtcaaccatg atctggacag atttgggagc 2880

tcggtataag ctctccacct atcttattct gtatagttta ggcttaaagt ttatccagga 2940

gatgttgctg aagtcgattt gagtccactt cctcactggt agctatacga ctttgatggt 3000

cgttgtaggg gctgtattag gtctcgatca aacacaaata gaattaaatg gtactcgagt 3060

ccactgaagg tggcttctcc gtcttccgta gccgtgccga aatccttaca gcttggtgtg 3120

tgtgactttt ggttacgccg tctgactttt gtggtgagct aactagagat catgctatat 3180

ctcctgattt aatacaatgc tcatacataac attccacctg gaactgctag caacgtttga 3240

cttgcaattgt gcaacgccct ttgcagagct atcgatgat caatagtgcc acgttctaaa 3300

ttcaaccaac gcaggtgccc caagccttcg acatccggat gtatttcgaa aacctcatgg 3360

cgattgcagt cctcagattc atgttcattc caatgctcat tggagaataa aagggtcaca 3420

gggaataagt tcaaactcga gatacttgag aatattgaaa gccaaaggac cctctatgct 3480

ccaagctaga gtctcagcct ggaaagcaaa tccaaatgaa gctatgctac ctccaattcc 3540

tcatacatctt atctataata cagagtcgaa gaatatcctc ttgacaccgc tccgtcctcc 3600

gacttcaata aggagcttac tctccttga caccaccctt ccagttcttc tcggcggttct 3660

ggagggaggc cttgtcggtc ttgggctggc cctggctgag aaagctgttg gcagccttaa 3720

agggacgctg gaggtcacca gtcgctggct tcccgaagac gtggatctta accagattcg 3780

aaagcgctt cagcgatga tcgactggat cagaagagcg ttggtgtact tgaagtacag 3840

atgcatgacg gccatcatgc caacgcccat gaactggctc ttaatgagct ggcggaactg 3900

gcccttatcg tactccatgt tggtagttgt gacaggacga ggctcctcgc cgcttccaag 3960

cggagcaggc tcgacgtatt tcagtgtcga aagatctgat caagagacag gatgaggatc 4020

gtttcgcattg attgaacaag atggattgca cgcaggttct ccggccgctt ggggtggagag 4080

gctattcggc tatgactggg cacaacagac aatcggtgc tctgatgccg ccgtgttccg 4140

gctgtcagcg cagggggcgc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa 4200

tgaactgcag gacgaggcag cgcggctatc gtggctggcc acgacgggcg ttccttgccg 4260

agctgtgctc gacgttgtca ctgaagcggg aagggaactg ctgctattgg gcgaagtgcc 4320

ggggcaggat ctctgtcat ctacacttgc tctgcccag aaagtatcca tcatggctga 4380

tgcaatgcgg cggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa 4440

acatcgcac gagcgagcac gtactcggat ggaagccggt cttgtcgatc aggatgatct 4500

ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcat 4560

ccccgacggc gaggatctcg tcgtgacca tggcgatgcc tgcttgccga atatcatggt 4620

ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta 4680

tcaggacata gcgttggcta cccgtgatat tgctgaagag cttggcggcg aatgggctga 4740

ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcacgc ccttctatcg 4800

ccttcttgac gagttcttct gagcgggact ctggggttcg aaatgaccga ccaagcgacg 4860

cccaacctgc catcacgaga ttctgattcc accgcgcct tctatgaaag gttgggcttc 4920

ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct catgctggag 4980

ttcttcgccc accccgggct cgatccctc gcgagttggt tcagctgctg cctgaggctg 5040

gacgacctcg cggagttcta ccggcagtgc aaatccgtcg gcatccagga aaccagcagc 5100

ggctatccgc gcatccatgc ccccgaactg caggagtggg gaggcacgat ggccgctttg 5160

gtccggatct ttgtgaagga accttacttc tgtgggtgtga cataattgga caaactacct 5220

acagagattt aaagctctaa ggtaaataa aaatttttta gtgtataatg tgttaaacta 5280

ctgattctaa ttgtttgtgt attttagatt ccaacctatg gaactgatga atgggagcag 5340

tggtggaatg cctttaatga ggaaaacctg ttttgctcag aagaaatgcc atctagtgat 5400

gatgaggcta ctgctgactc tcaacattct actcctccaa aaaagaagag aaaggtagaa 5460

gacccaagc actttccttc agaattgcta agttttttga gtcatgctgt gtttagtaat 5520

agaactcttg cttgctttgc tatttacacc acaaaggaaa aagctgcact gctatacaag 5580

aaaattatgg aaaaatattc tgtaaccttt ataagtaggc ataacagtta taatcataac 5640

atactgtttt ttcttactcc acacagggcat agagtgtctg ctattaataa ctatgctcaa 5700

aaattgtgta ccttttagctt tttaatttgt aaaggggtta ataaggaata tttgatgtat 5760

agtccttga ctagagatca taatcagcca taccacattt gtagaggttt tacttgcttt 5820

aaaaaacctc ccacacctcc ccctgaacct gaaacataaa atgaatgcaa ttgttggtgt 5880

taacttgttt attgcagctt ataatggta caaataaagc aatagcatca caaatctcac 5940

aaataaagca tttttttcac tgcattctag ttgtggtttg tccaaactca tcaatgtatc 6000

ttatcatgtc tggatctgac ggggtgcgat gatcgtgctc ctgctggtga ggacccggct 6060

aggctggcgg ggttgcctta ctggtagca gaatgaatca ccgatacgcg agcgaacgtg 6120

aagcgactgc tgctgcaaaa cgtctgcgac ctgagcaaca acatgaatgg tcttcggttt 6180

ccgtgtttcg taaagtctgg aaacgcggaa gtcagcgctc ttccgcttcc tcgctcactg 6240

actcgtgcg ctcggtcgtt cggctgcggc gagcggatc agctcactca aaggcggtaa 6300

tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 6360

aaagggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt tttccatag 6420

gctccgcccc cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaacct 6480

gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 6540

tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 6600

ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 6660

ctgtgtgcac gaaccccccg ttcagcccgga ccgctgcgcc ttatccggta actatcgtct 6720

tgagtccaac ccggtgaagac acgacttatc gccactggca gcagccactg gtaacaggat 6780

tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtgggtggc ctaactacgg 6840

ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 6900

aagagttggt agctcttgat cgggcaaaca aaccaccgct ggtagcgggtg gtttttttgt 6960

ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 7020

tacgggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 7080

atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 7140

aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 7200

ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac 7260

tacgatacgg gagggcttac catctggccc cagtgcctga atgataccgc gagaccacg 7320

ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag 7380

tggtcctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt 7440

aagtagttcg ccagttaata gtttgcgcaa cgttggtgcc attgctgcag gcatcgtggt 7500

gtcacgctcg tcgtttggta tggcttcatt cagctccggt tccaacgat caaggcgagt 7560

tacatgatcc cccatgttgt gcaaaaaagc ggtagctcc ttcggtcctc cgatcgttgt 7620

cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct 7680

tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt 7740

ctgagaatag tgtatgcggc gaccgagttg ctcttgcccc gcgtcaacac gggataatac 7800

cgcgccacat agcagaactt taaaagtgt catcattgga aaacgttctt cggggcgaaa 7860

actctcaagg atcttaccgc tgttgagatc cagttcgatg taaccactc gtgcacccaa 7920

ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 7980

aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct 8040

ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 8100

atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 8160

tgacgtctaa gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag 8220

gccctttcgt cttcaagaat tcgcggccgc aattaaccct cactaaagga tccctatagt 8280

gagtcgtatt atgcggccgc gaattctcat gtttgaccgc ttatcatcga taagctctgc 8340

tttttgttga cttccattgt tcattccacg gacaaaaaca gagaaaggaa acgacagagg 8400

ccaaaaagct cgctttcagc acctgtcgtt tcctttcttt tcagagggtta ttttaaataa 8460

aaacattaag ttatgacgaa gaagaacgga aacgccttaa accggaaaat tttcataaat 8520

agcgaaaacc cgcgagggtcg ccgccccgta acaaggcgga tcgccggaaa ggacccgcaa 8580

atgataataa ttatcaattg catactatcg acggcactgc tgccagataa caccaccggg 8640

gaaacattcc atcatgatgg ccgtgctggac ataggaagcc agttcatcca tcgctttctt 8700

gtctgctgcc atttgctttg tgacatccag cgccgcacat tcagcagcgt ttttcagcgc 8760

gttttcgata aacgtttcaa tgttgggtatc aacaccagggt ttaactttga acttatcggc 8820

actgacgggtt accttgttct gcgctggctc atcacgcagg ataccaaggc tgatgttgta 8880

gatattggtc accggctgag ggttttcgat tgccgctgcg tggatagcac catttgcgat 8940

caggcngtcc ttgatgaatg aactccatt gcgaataagt tcgaaggaga cgggtgtcacg 9000

aatgcgctgg tccagctcgg tcgattgcct tttgtgcagc agaggtatca atctcaacgc 9060

caaggctcat cgaagcgcaa tattgctgct caccaaaacg cgtattgacc aggtgttcaa 9120

cggcaaattt ctgcccttct gatgtcagaa aggcaaagtg attttctttc tggatttcag 9180

ttgctgtgtg tcggtttcag caaaaccaag ctgcgcgaat tcggctgtgc agatttagaa 9240

ggcagatcac cagacagcaa cggccaacgg aaaacagcgc atacagaaca tccgtcgccg 9300

cgccgacaac gtgataattt ttatgacca tgatttattt ctttttagac gtgagcctgt 9360

cgcacagcaa agccgccgaa agttcctcga agctagcttc agacgtgtct agatacgtct 9420

gctttttgtt gacttccatt gttcattcca cggacaaaaa cagagaaagg aaacgacaga 9480

ggccaaaaag ctcgctttca gcacctgtcg tttcctttct tttcagaggg tattttaaat 9540

aaaaacatta agttatgacg aagaagaacg gaaacgcctt aaaccgaaa attttcataa 9600

atagcgaaaa cccgcgaggt cgccgccccg taacaaggcg gatcgccgga aaggacccgc 9660

aaatgataat aattatcaat tgcatactat cgacggcact gctgccagat aacaccaccg 9720

gggaaacatt ccatcatgat ggccgtgcgg acataggaag ccagttcatc catcgctttc 9780

ttgtctgctg ccatttgctt tgtgacatcc agcgccgcac attcagcagc gtttttcagc 9840

gcgttttcga tcaacgtttc aatgttggtg tcaacaccag gtttaacttt gaacttatcg 9900

gcactgacgg ttaccttggt ctgcgctggc tcatcacgca ggataccaag gctgatgttg 9960

tagatatggg tcaccggctg agggttttcg attgccgctg cgtggatagc accatttgcg 10020

atcaggcngt ccttgatgaa tgacactcca ttgcgaataa gttcgaagga gacgggtgtca 10080

cgaatgcgtt ggtccagctc ggtcgattgc cttttgtgca gcagaggtat caatctcaac 10140

gccaaggctc atcgaagcgc aatattgctg ctcacaaaaa cgcgtattga ccagggtgtc 10200

aacggcaaatt ttctgccctt ctgatgtcag aaaggcaaag tgattttctt tctgggtattc 10260

agttgctgtg tgtcggtttc agcaaaacca agctcgcgca attcggctgt gcagatttag 10320

aaggcagatc accagacagc aacggccaac ggaaaacagc gcatacagaa catccgtcgc 10380

cgcgccgaca acgtgataat ttttatgacc catgatttat ttccttttag acgtgagcct 10440

gtcgcacagc aaagccgccg aaagtctctc gaccgatgcc cttgagagcc ttcaaccag 10500

tcagctcctt ccggtgggcg cggggcatga ctatcgtcgc cgcacttatg actgtcttct 10560

ttatcatgca actcgtagga caggtgccgg cagcgctctg ggtcattttc ggcgaggacc 10620

gctttcgctg gagcgcgacg atgatcggcc tgtcgcttgc ggtattcgga atcttgcacg 10680

ccctcgctca agccttcgtc actggtcccg ccaccaaacg tttcggcgag aagcaggcca 10740
ttatcgccgg catggcggcc gacgcgctgg gctacgtctt gctggcgttc gcgacgcgag 10800
gctggatggc cttccccatt atgattcttc tcgcttcggg cggcatcggg atgcccgcgt 10860
tgcaggccat gctgtccagg caggtagatg acgaccatca gggacagctt caaggatcgc 10920
tcgcggctct taccagccta acttcgatca ttggaccgct gatcgtcacg gcgatttatg 10980
ccgcctcggc gagcacatgg aacgggttgg catggattgt aggcgccgcc ctataccttg 11040
tctgcctccc cgcgttgctg cgcggtgcat ggagccgggc cacctcgacc tgaatggaag 11100
ccggcggcac ctcgctaacg gattcaccac tccaagaatt ggagccaatc aattcttgcg 11160
gagaactgtg aatgcgcaaa ccaacccttg gcagaacata tccatcgctt ccgccatctc 11220
cagcagccgc acgcggcgca tctcgggcag cgttgggtcc tgcagatccg gctgtggaat 11280
gtgtgtcagt tagggtgtgg aaagtcccca ggctccccag caggcagaag tatgcaaagc 11340
atgcattctca attagtcagc aaccaggtgt ggaaagtccc caggctcccc agcaggcaga 11400
agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgccctt aactccgcc 11460
atcccccccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt 11520
tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga 11580
ggcttttttg gaggcctagg cttttgcaaa a 11611

<210> 5

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 5

cgatgtagga gggcgtggat a

21

<210> 6

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 6

gcttctgcgg gcgatttg t

21

<210> 7

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 7

tgagaatatc accggaattg

20

<210> 8

<211> 21

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 8

agctcgacat actgttcttc c

21

<210> 9

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 9

gtgaatggaa atcccatcgc tgtc

24

<210> 10

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 10

agtgggtact ctaaaggcca tacc

24

<210> 11

<211> 1771

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (166)..(1155)

<400> 11

ggcacgagct tgcacgcaag tcagcgcgcg caagtcaaca cctgccgggc cacagcctca

60

aataataaag agctcaagcg tttgtgcgcc tcgacgtggc cagtctgcac tgccttgaac

120

ccgcgagtct cccgccgcac tgactgccat agcacagcta gacga atg cag cta gca

177

Met Gln Leu Ala

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225
 Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys
 5 10 15 20

gag aag gag aag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273
 Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp
 25 30 35

gcg acc cag tac tcg ctt ccg tca gaa gag tca gac gcg gcc cgc ccg 321
 Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro
 40 45 50

gga ctg aag aat gcc tac aag cca cca cct tcc gac aca aag ggc atc 369
 Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile
 55 60 65

aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac 417
 Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His
 70 75 80

gcc att ttt caa atc aag ctt ccg acc tcc ttg gac cag ctg cac tgg 465
 Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp
 85 90 95 100

ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc 513
 Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser Gly Thr Ser Ser
 105 110 115

ctg ctc gac atc gtc gta gta ttc ttt gtc ctg gag ttc ctg tac aca 561
 Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu Phe Leu Tyr Thr
 120 125 130

ggc ctt ttt atc acc acg cat gat gct atg cat ggc acc atc gcc atg 609
 Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly Thr Ile Ala Met
 135 140 145

aga aac agg cag ctt aat gac ttc ttg ggc aga gta tgc atc tcc ttg 657
 Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val Cys Ile Ser Leu
 150 155 160

tac gcc tgg ttt gat tac aac atg ctg cac cgc aag cat tgg gag cac	705
Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys His Trp Glu His	
165 170 175 180	
cac aac cac act ggc gag gtg ggc aag gac cct gac ttc cac agg gga	753
His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp Phe His Arg Gly	
185 190 195	
aac cct ggc att gtg ccc tgg ttt gcc agc ttc atg tcc agc tac atg	801
Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met	
200 205 210	
tcg atg tgg cag ttt gcg cgc ctc gca tgg tgg acg gtg gtc atg cag	849
Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr Val Val Met Gln	
215 220 225	
ctg ctg ggt gcg cca atg gcg aac ctg ctg gtg ttc atg gcg gcc gcg	897
Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala	
230 235 240	
ccc atc ctg tcc gcc ttc cgc ttg ttc tac ttt ggc acg tac atg ccc	945
Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Met Pro	
245 250 255 260	
cac aag cct gag cct ggc gcc gcg tca ggc tct tca cca gcc gtc atg	993
His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser Pro Ala Val Met	
265 270 275	
aac tgg tgg aag tcg cgc act agc cag gcg tcc gac ctg gtc agc ttt	1041
Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp Leu Val Ser Phe	
280 285 290	
ctg acc tgc tac cac ttc gac ctg cac tgg gag cac cac cgc tgg ccc	1089
Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His His Arg Trp Pro	
295 300 305	
ttc gcc ccc tgg tgg gag ctg ccc aac tgc cgc cgc ctg tct ggc cga	1137
Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg Leu Ser Gly Arg	
310 315 320	
ggt ctg gtt cct gcc tag ctggacacac tgcagtgggc cctgctgcc	1185

Gly Leu Val Pro Ala

325

gctgggcatg caggttgtgg caggactggg tgaggtgaaa agctgcaggc gctgctgccg 1245
 gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggaggggg tttgtagctg 1305
 tcgagcttgc cccatggatg aagctgtgta gtggtgcagg gagtacaccc acaggccaac 1365
 acccttgcag gagatgtctt gcgtcgggag gagtgttggg cagtgtagat gctatgattg 1425
 tatcttaatg ctgaagcctt taggggagcg acacttagtg ctgggcaggc aacgccctgc 1485
 aaggtgcagg cacaagctag gctggacgag gactcgggtg caggcaggtg aagaggtgcg 1545
 ggaggggtgt gccacacca ctgggcaaga ccatgctgca atgctggcgg tgtggcagtg 1605
 agagctgcgt gattaactgg gctatggatt gtttgagcag tctcacttat tctttgatat 1665
 agatactggg caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgctgc 1725
 ccctgcgctt atgaagctgt aacaataaag tgggtcaaaa aaaaaa 1771

<210> 12

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 12

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala

1

5

10

15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val

20

25

30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp
50 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala
65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp
85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser
100 105 110

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val
145 150 155 160

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp
180 185 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr
210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser
260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp
275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His
290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg
305 310 315 320

Leu Ser Gly Arg Gly Leu Val Pro Ala
325

<210> 13

<211> 1662

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (168) .. (1130)

<400> 13

```

cggggcaact caagaaattc aacagctgca agcgcgcccc agcctcacag cgccaagtga      60

gctatcgacg tggttgtgag cgctcgacgt ggtccactga cgggcctgtg agcctctgcg      120

ctccgtcctc tgccaaatct cgcgtcgggg cctgcctaag tcgaaga atg cac gtc      176
                                     Met His Val
                                     1

gca tcg gca cta atg gtc gag cag aaa ggc agt gag gca gct gct tcc      224
Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala Ala Ala Ser
      5              10              15

agc cca gac gtc ttg aga gcg tgg gcg aca cag tat cac atg cca tcc      272
Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His Met Pro Ser
20              25              30              35

gag tcg tca gac gca gct cgt cct gcg cta aag cac gcc tac aaa cct      320
Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala Tyr Lys Pro
              40              45              50

cca gca tct gac gcc aag ggc atc acg atg gcg ctg acc atc att ggc      368
Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr Ile Ile Gly
              55              60              65

acc tgg acc gca gtg ttt tta cac gca ata ttt caa atc agg cta ccg      416
Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile Arg Leu Pro
      70              75              80

aca tcc atg gac cag ctt cac tgg ttg cct gtg tcc gaa gcc aca gcc      464
Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu Ala Thr Ala
      85              90              95

cag ctt ttg ggc gga agc agc agc cta ctg cac atc gct gca gtc ttc      512
Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala Ala Val Phe
100              105              110              115

att gta ctt gag ttc ctg tac act ggt cta ttc atc acc aca cat gac      560
Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp
              120              125              130

gca atg cat ggc acc ata gct ttg agg cac agg cag ctc aat gat ctc      608

```

Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu	
135 140 145	
ctt ggc aac atc tgc ata tca ctg tac gcc tgg ttt gac tac agc atg	656
Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Ser Met	
150 155 160	
ctg cat cgc aag cac tgg gag cac cac aac cat act ggc gaa gtg ggg	704
Leu His Arg Lys His Trp Glu His His Asn His Thr Gly Glu Val Gly	
165 170 175	
aaa gac cct gac ttc cac aag gga aat ccc ggc ctt gtc ccc tgg ttc	752
Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val Pro Trp Phe	
180 185 190 195	
gcc agc ttc atg tcc agc tac atg tcc ctg tgg cag ttt gcc cgg ctg	800
Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe Ala Arg Leu	
200 205 210	
gca tgg tgg gca gtg gtg atg caa atg ctg ggg gcg ccc atg gca aat	848
Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro Met Ala Asn	
215 220 225	
ctc cta gtc ttc atg gct gca gcc cca atc ttg tca gca ttc cgc ctc	896
Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu	
230 235 240	
ttc tac ttc ggc act tac ctg cca cac aag cct gag cca ggc cct gca	944
Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro Gly Pro Ala	
245 250 255	
gca ggc tct cag gtg atg gcc tgg ttc agg gcc aag aca agt gag gca	992
Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr Ser Glu Ala	
260 265 270 275	
tct gat gtg atg agt ttc ctg aca tgc tac cac ttt gac ctg cac tgg	1040
Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp	
280 285 290	
gag cac cac agg tgg ccc ttt gcc ccc tgg tgg cag ctg ccc cac tgc	1088
Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu Pro His Cys	
295 300 305	

cgc cgc ctg tcc ggg cgt ggc ctg gtg cct gcc ttg gca tga 1130
Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala
310 315 320

cctgggtccct ccgctggtga ccagcgtct gcacaagagt gtcattgctac aggggtgctgc 1190

ggccagtggc agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccactgagca 1250

ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg 1310

ggcgtgctac tgacaatggg cgtgctactg gggctctggca gtggctagga tggagtttga 1370

tgcattcagt agcgggtggc aacgtcatgt ggatgggtgga agtgctgagg ggtttaggca 1430

gccggcattt gagagggcta agttataaat cgcattgctgc tcatgcgcac atatctgcac 1490

acagccaggg aaatcccttc gagagtgatt atgggacact tgtattgggt tcgtgctatt 1550

gttttattca gcagcagtac ttagtgaggg tgagagcagg gtggtgagag tggagtgagt 1610

gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct 1662

<210> 14

<211> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala
1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His
20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala
35 40 45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu
85 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala
100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
115 120 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu
130 135 140

Asn Asp Leu Leu Gly Asn Ile Cys Ile Ser Leu Tyr Ala Trp Phe Asp
145 150 155 160

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly
165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val
180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe
195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro

210	215	220	
Met Ala Asn Leu Leu Val Phe Met Ala Ala Ala Pro Ile Leu Ser Ala			
225	230	235	240
Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro			
	245	250	255
Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr			
	260	265	270
Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp			
	275	280	285
Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu			
	290	295	300
Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala			
305	310	315	320

<210> 15
 <211> 729
 <212> DNA
 <213> Agrobacterium aurantiacum

<220>
 <221> CDS
 <222> (1)..(729)

<400> 15
 atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gct tgg ctg gcc ctg cat gtg cat	96
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His	
20 25 30	
gcg ctg tgg ttt ctg gac gca gcg gcg cat ccc atc ctg gcg atc gca	144
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala	
35 40 45	
aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg	192
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	
50 55 60	
cat gac gcg atg cac ggg tcg gtg gtg ccg ggg cgt ccg cgc gcc aat	240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	
65 70 75 80	
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg	288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	
85 90 95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc	336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	
100 105 110	
gac gac gac ccc gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc	384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	
115 120 125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc	432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	
130 135 140	
gtc atc gtg acg gtc tat gcg ctg atc ctt ggg gat cgc tgg atg tac	480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr	
145 150 155 160	
gtg gtc ttc tgg ccg ctg ccg tcg atc ctg gcg tcg atc cag ctg ttc	528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	
165 170 175	
gtg ttc ggc acc tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg	576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	

180	185	190	
gac cgc cac aat gcg cgg tcg tcg cgg atc agc gac ccc gtg tcg ctg			624
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu			
195	200	205	
ctg acc tgc ttt cac ttt ggc ggt tat cat cac gaa cac cac ctg cac			672
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His			
210	215	220	
ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac			720
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp			
225	230	235	240
acc gca tga			729
Thr Ala			

<210> 16

<211> 242

<212> PRT

<213> Agrobacterium aurantiacum

<400> 16

Met	Ser	Ala	His	Ala	Leu	Pro	Lys	Ala	Asp	Leu	Thr	Ala	Thr	Ser	Leu
1				5					10					15	

Ile	Val	Ser	Gly	Gly	Ile	Ile	Ala	Ala	Trp	Leu	Ala	Leu	His	Val	His
			20					25					30		

Ala	Leu	Trp	Phe	Leu	Asp	Ala	Ala	Ala	His	Pro	Ile	Leu	Ala	Ile	Ala
			35					40					45		

Asn	Phe	Leu	Gly	Leu	Thr	Trp	Leu	Ser	Val	Gly	Leu	Phe	Ile	Ile	Ala
			50				55				60				

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

Thr Ala

<210> 17

<211> 1631

<212> DNA

<213> Alcaligenes sp.

<220>

<221> CDS

<222> (99)..(827)

<400> 17

ctgcaggccg ggcccgggtgg ccaatgggtcg caaccggcag gactggaaca ggacggcggg 60

ccgggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116
 Met Ser Gly Arg Lys Pro
 1 5

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164
 Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile
 10 15 20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212
 Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp
 25 30 35

gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260
 Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr
 40 45 50

tgg ctg tcg gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308
 Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly
 55 60 65 70

tcc gtg gtg ccg ggg cgg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356

Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ala Ile Gly Gln Leu

75	80	85	
gcg ctg tgg ctc tat gcg ggg ttc tcg tgg ccc aag ctg atc gcc aag			404
Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys			
90	95	100	
cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc			452
His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe			
105	110	115	
ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat			500
Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr			
120	125	130	
ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat			548
Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro Val Ile Val Thr Thr Tyr			
135	140	145	150
gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc			596
Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val			
155	160	165	
ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg			644
Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu			
170	175	180	
ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gcg agg			692
Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg			
185	190	195	
tcg acc ggc atc ggc gac ccg ttg tca cta ctg acc tgc ttc cat ttc			740
Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe			
200	205	210	
ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg			788
Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp Trp			
215	220	225	230
cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct			837
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala			
235	240		

cattgtcgtg gcgacagtcc tcgtgatgga gctgaccgcc tattccgtcc accgctggat 897
tatgcacggc cccctaggct ggggctggca caagtcccat cacgaagagc acgaccacgc 957
gttggaagaag aacgacctct acggcgtcgt cttcgcgggtg ctggcgacga tcctcttcac 1017
cgtggggcgcc tattggtggc cgggtgctgtg gtggatcgcc ctgggcatga cggctctatgg 1077
gttgatctat ttcatactgc acgacgggct tgtgcatcaa cgctggccgt ttcggtatat 1137
tccgcggcgg ggctatttcc gcaggctcta ccaagctcat cgcctgcacc acgcggtcga 1197
ggggcgggac cactgcgtca gcttcggctt catctatgcc ccacccgtgg acaagctgaa 1257
gcaggatctg aagcggtcgg gtgtcctgcg ccccaggac gagcgtccgt cgtgatctct 1317
gatcccgggc tggccgcatg aaatccgacg tgctgctggc aggggcccgc cttgccaacg 1377
gactgatcgc gctggcgatc cgcaaggcgc ggcccgaact tcgcgtgctg ctgctggacc 1437
gtgcggcggg cgcctcggac gggcatactt ggtcctgcca cgacaccgat ttggcgccgc 1497
actggctgga ccgcctgaag ccgatcagc gtggcgactg gcccgatcag gaggtgcggt 1557
tcccagacca ttcgcgaagg ctccgggccg gatatggctc gatcgacggg cgggggctga 1617
tgcggtgcgt gacc 1631

<210> 18

<211> 242

<212> PRT

<213> *Alcaligenes* sp.

<400> 18

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu

1

5

10

15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe
20 25 30

Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu
35 40 45

Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr
100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly
115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe
165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro
180 185 190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu
 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly
 225 230 235 240

Arg Ala

<210> 19
 <211> 729
 <212> DNA
 <213> *Paracoccus marcusii*

<220>
 <221> CDS
 <222> (1)..(729)

<400> 19
 atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc aca agc ctg 48
 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
 1 5 10 15
 atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat 96
 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
 20 25 30
 gcg ctg tgg ttt ctg gac gcg gcg gcc cat ccc atc ctg gcg gtc gcg 144
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
 35 40 45
 aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala	
50 55 60	
cat gac gcg atg cac ggg tcg gtc gtg ccg ggg cgt ccg cgc gcc aat	240
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn	
65 70 75 80	
gcg gcg atg ggc cag ctt gtc ctg tgg ctg tat gcc gga ttt tcg tgg	288
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp	
85 90 95	
cgc aag atg atc gtc aag cac atg gcc cat cac cgc cat gcc gga acc	336
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr	
100 105 110	
gac gac gac cca gat ttc gac cat ggc ggc ccg gtc cgc tgg tac gcc	384
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala	
115 120 125	
cgc ttc atc ggc acc tat ttc ggc tgg cgc gag ggg ctg ctg ctg ccc	432
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro	
130 135 140	
gtc atc gtg acg gtc tat gcg ctg atc ctg ggg gat cgc tgg atg tac	480
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr	
145 150 155 160	
gtg gtc ttc tgg ccg ttg ccg tcg atc ctg gcg tcg atc cag ctg ttc	528
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe	
165 170 175	
gtg ttc ggc act tgg ctg ccg cac cgc ccc ggc cac gac gcg ttc ccg	576
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro	
180 185 190	
gac cgc cat aat gcg cgg tcg tcg cgg atc agc gac cct gtg tcg ctg	624
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu	
195 200 205	
ctg acc tgc ttt cat ttt ggc ggt tat cat cac gaa cac cac ctg cac	672
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His	
210 215 220	

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
225 230 235 240

acc gca tga 729
Thr Ala

<210> 20

<211> 242

<212> PRT

<213> *Paracoccus marcusii*

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala
35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr

50/358

100	105	110
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala 115 120 125		
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro 130 135 140		
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr 145 150 155 160		
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 165 170 175		
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro 180 185 190		
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu 195 200 205		
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 215 220		
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 240		
Thr Ala		

<210> 21

<211> 1629

<212> DNA

<213> Synechocystis sp.

<220>

<221> CDS

<222> (1)..(1629)

<400> 21

atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta 48
Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
1 5 10 15

gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta 96
Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
20 25 30

gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg 144
Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met
35 40 45

ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac 192
Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
50 55 60

gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag 240
Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln
65 70 75 80

tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg 288
Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
85 90 95

ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt 336
Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
100 105 110

gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa 384
Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt 432

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe	
130 135 140	
aat gct ccg ccc cag gct tta cta gat tta gcc ctg aac tat ggt tgg	480
Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp	
145 150 155 160	
gaa aac tta aaa tcc gtg ctg gcg atc gcc ggg tcg aaa acc aag gcg	528
Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala	
165 170 175	
ttg gat ttt atc cgc act atg atc ggc tcc ccg gaa gat gtg ctc aat	576
Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn	
180 185 190	
gaa tgg ttc gac agc gaa cgg gtt aaa gct cct tta gct aga cta tgt	624
Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys	
195 200 205	
tcg gaa att ggc gct ccc cca tcc caa aag ggt agt agc tcc ggc atg	672
Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met	
210 215 220	
atg atg gtg gcc atg cgg cat ttg gag gga att gcc aga cca aaa gga	720
Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly	
225 230 235 240	
ggc act gga gcc ctc aca gaa gcc ttg gtg aag tta gtg caa gcc caa	768
Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln	
245 250 255	
ggg gga aaa atc ctc act gac caa acc gtc aaa cgg gta ttg gtg gaa	816
Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu	
260 265 270	
aac aac cag gcg atc ggg gtg gag gta gct aac gga gaa cag tac cgg	864
Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg	
275 280 285	
gcc aaa aaa ggc gtg att tct aac atc gat gcc cgc cgt tta ttt ttg	912
Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu	
290 295 300	

54/358

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
 465 470 475 480

tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488
 Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
 485 490 495

ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536
 Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
 500 505 510

ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584
 Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
 515 520 525

aat tgc gct cgg gtc ttt tta aaa caa caa cgt cgt ttt tgg taa 1629
 Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
 530 535 540

<210> 22

<211> 542

<212> PRT

<213> Synechocystis sp.

<400> 22

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
 1 5 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu
 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met
 35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His
 50 55 60

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln
65 70 75 80

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly
85 90 95

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys
100 105 110

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln
115 120 125

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe
130 135 140

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp
145 150 155 160

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala
165 170 175

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn
180 185 190

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys
195 200 205

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met
210 215 220

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly

225					230						235					240
Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln																
				245						250					255	
Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu				260					265					270		
Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg				275				280					285			
Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu				290				295				300				
Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly	305						310				315					320
Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Asn Glu Ala Ile Leu Lys				325						330				335		
Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly			340						345				350			
Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His		355					360					365				
Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala	370					375					380					
Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met	385				390					395						400

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr
405 410 415

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr
420 425 430

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr
435 440 445

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu
450 455 460

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val
465 470 475 480

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu
485 490 495

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
500 505 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg
515 520 525

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
530 535 540

<210> 23

<211> 776

<212> DNA

<213> Bradyrhizobium sp.

<220>

<221> CDS

<222> (1) . . (774)

<400> 23

atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cag cgc 48

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg

1 5 10 15

gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc 96

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile

20 25 30

atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg ccg 144

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro

35 40 45

ctg acc ctt cac agc ctg ctg ccg gct ttg cct ctg gtg gtg ctg caa 192

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln

50 55 60

acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac 240

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His

65 70 75 80

ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga caa 288

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln

85 90 95

ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc 336

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val

100 105 110

gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat 384

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp

115

ttc gac gag gtg ccg ccg cac ggc ttc tgg cac tgg ttc gcc agc ttt 432

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe

130

ttc ctg cac tat ttc ggc tgg aag cag gtc gcg atc atc gca gcc gtc 480
 Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
 145 150 155 160
 tcg ctg gtt tat cag ctc gtc ttc gcc gtt ccc ttg cag aac atc ctg 528
 Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
 165 170 175
 ctg ttc tgg gcg ctg ccc ggg ctg ctg tcg gcg ctg cag ctg ttc acc 576
 Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
 180 185 190
 ttc ggc acc tat ctg ccg cac aag ccg gcc acg cag ccc ttc gcc gat 624
 Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
 195 200 205
 cgc cac aac gcg cgg acg agc gaa ttt ccc gcg tgg ctg tcg ctg ctg 672
 Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
 210 215 220
 acc tgc ttc cac ttc ggc ttt cat cac gag cat cat ctg cat ccc gat 720
 Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
 225 230 235 240
 gcg ccg tgg tgg cgg ctg ccg gag atc aag cgg cgg gcc ctg gaa agg 768
 Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255
 cgt gac ta 776
 Arg Asp

<210> 24

<211> 258

<212> PRT

<213> Bradyrhizobium sp.

<400> 24

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg

60/358

1	5	10	15
Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile	20	25	30
Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro	35	40	45
Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln	50	55	60
Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His	65	70	75
Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln	85	90	95
Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val	100	105	110
Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp	115	120	125
Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe	130	135	140
Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val	145	150	155
Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu	165	170	175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
 180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
 195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
 210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
 225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255

Arg Asp

<210> 25

<211> 777

<212> DNA

<213> Nostoc sp.

<220>

<221> CDS

<222> (1)..(777)

<400> 25

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta 48
 Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
 1 5 10 15

ttg tca tcg aca atc aga gat gat aaa aat att aat aag ggt ata ttt 96
 Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe

62/358

20	25	30	
att gcc tgc ttt atc tta ttt tta tgg gca att agt tta atc tta tta			144
Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu			
35	40	45	
ctc tca ata gat aca tcc ata att cat aag agc tta tta ggt ata gcc			192
Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala			
50	55	60	
atg ctt tgg cag acc ttc tta tat aca ggt tta ttt att act gct cat			240
Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His			
65	70	75	80
gat gcc atg cac ggc gta gtt tat ccc aaa aat ccc aga ata aat aat			288
Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn			
85	90	95	
ttt ata ggt aag ctc act cta atc ttg tat gga cta ctc cct tat aaa			336
Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys			
100	105	110	
gat tta ttg aaa aaa cat tgg tta cac cac gga cat cct ggt act gat			384
Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp			
115	120	125	
tta gac cct gat tat tac aat ggt cat ccc caa aac ttc ttt ctt tgg			432
Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp			
130	135	140	
tat cta cat ttt atg aag tct tat tgg cga tgg acg caa att ttc gga			480
Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly			
145	150	155	160
tta gtg atg att ttt cat gga ctt aaa aat ctg gtg cat ata cca gaa			528
Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu			
165	170	175	
aat aat tta att ata ttt tgg atg ata cct tct att tta agt tca gta			576
Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val			
180	185	190	

caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt 624
 Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
 195 200 205

ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt 672
 Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
 210 215 220

tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac 720
 Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768
 Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

tct tta taa 777
 Ser Leu

<210> 26

<211> 258

<212> PRT

<213> Nostoc sp.

<400> 26

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu
 1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe
 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
 35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala
 50 55 60

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His
65 70 75 80

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn
85 90 95

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys
100 105 110

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp
115 120 125

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp
130 135 140

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly
145 150 155 160

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu
165 170 175

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val
180 185 190

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
195 200 205

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe
210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His
 225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
 245 250 255

Ser Leu

<210> 27

<211> 789

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(789)

<400> 27

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa 48
 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
 1 5 10 15

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96
 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
 20 25 30

att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144
 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
 35 40 45

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192
 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
 50 55 60

atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat 240
 Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
 65 70 75 80

ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca	288
Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser	
85 90 95	
cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag	336
Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys	
100 105 110	
aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat	384
Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp	
115 120 125	
ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc	432
Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe	
130 135 140	
atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta	480
Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu	
145 150 155 160	
ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc	528
Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile	
165 170 175	
tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat	576
Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr	
180 185 190	
ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat	624
Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr	
195 200 205	
ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc	672
Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile	
210 215 220	
gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat	720
Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His	
225 230 235 240	
gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac	768

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
245 250 255

aat tca gta acc aat tcg taa
Asn Ser Val Thr Asn Ser
260

789

<210> 28
<211> 262
<212> PRT
<213> Nostoc punctiforme

<400> 28

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
1 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
20 25 30

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
50 55 60

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
65 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
85 90 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
100 105 110

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
115 120 125

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
130 135 140

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
145 150 155 160

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
165 170 175

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
180 185 190

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
195 200 205

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
210 215 220

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
225 230 235 240

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
245 250 255

Asn Ser Val Thr Asn Ser
260

<211> 762
 <212> DNA
 <213> Nostoc punctiforme

<220>
 <221> CDS
 <222> (1)..(762)

<400> 29
 gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca 48
 Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
 1 5 10 15
 gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc 96
 Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
 20 25 30
 att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac 144
 Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp
 35 40 45
 atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa 192
 Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
 50 55 60
 aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat 240
 Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
 65 70 75 80
 ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca 288
 Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
 85 90 95
 ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa 336
 Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
 100 105 110
 aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat 384
 Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
 115 120 125

ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt 432
 Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
 130 135 140

atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att 480
 Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
 145 150 155 160

tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act 528
 Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
 165 170 175

tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat 576
 Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
 180 185 190

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag 624
 Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc 672
 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat 720
 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag 762
 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
 245 250

<210> 30

<211> 253

<212> PRT

<213> Nostoc punctiforme

<400> 30

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
 1 5 10 15

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
20 25 30

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp
35 40 45

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
50 55 60

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
65 70 75 80

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
85 90 95

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
100 105 110

Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
115 120 125

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
130 135 140

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
145 150 155 160

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
 180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
 195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
 210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
 225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
 245 250

<210> 31

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3) .. (971)

<400> 31

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc 47
 Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile
 1 5 10 15

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg 95
 Gly Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu
 20 25 30

tcg aag ctg cag tca atc agc gtc aag gcc cgc cgc gtt gaa cta gcc 143
 Ser Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala
 35 40 45

cgc gac atc acg cgg ccc aaa gtc tgc ctg cat gct cag cgg tgc tcg	191
Arg Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser	
50 55 60	
tta gtt cgg ctg cga gtg gca gca cca cag aca gag gag gcg ctg gga	239
Leu Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly	
65 70 75	
acc gtg cag gct gcc ggc gcg ggc gat gag cac agc gcc gat gta gca	287
Thr Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala	
80 85 90 95	
ctc cag cag ctt gac cgg gct atc gca gag cgt cgt gcc cgg cgc aaa	335
Leu Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys	
100 105 110	
cgg gag cag ctg tca tac cag gct gcc gcc att gca gca tca att ggc	383
Arg Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly	
115 120 125	
gtg tca ggc att gcc atc ttc gcc acc tac ctg aga ttt gcc atg cac	431
Val Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His	
130 135 140	
atg acc gtg ggc ggc gca gtg cca tgg ggt gaa gtg gct ggc act ctc	479
Met Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu	
145 150 155	
ctc ttg gtg gtt ggt ggc gcg ctc ggc atg gag atg tat gcc cgc tat	527
Leu Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr	
160 165 170 175	
gca cac aaa gcc atc tgg cat gag tcg cct ctg ggc tgg ctg ctg cac	575
Ala His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His	
180 185 190	
aag agc cac cac aca cct cgc act gga ccc ttt gaa gcc aac gac ttg	623
Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu	
195 200 205	
ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc	671

Phe	Ala	Ile	Ile	Asn	Gly	Leu	Pro	Ala	Met	Leu	Leu	Cys	Thr	Phe	Gly			
						210							215				220	
ttc	tgg	ctg	ccc	aac	gtc	ctg	ggg	gcg	gcc	tgc	ttt	gga	gcg	ggg	ctg	719		
Phe	Trp	Leu	Pro	Asn	Val	Leu	Gly	Ala	Ala	Cys	Phe	Gly	Ala	Gly	Leu			
						225							230				235	
ggc	atc	acg	cta	tac	ggc	atg	gca	tat	atg	ttt	gta	cac	gat	ggc	ctg	767		
Gly	Ile	Thr	Leu	Tyr	Gly	Met	Ala	Tyr	Met	Phe	Val	His	Asp	Gly	Leu			
						240							245				250	255
gtg	cac	agg	cgc	ttt	ccc	acc	ggg	ccc	atc	gct	ggc	ctg	ccc	tac	atg	815		
Val	His	Arg	Arg	Phe	Pro	Thr	Gly	Pro	Ile	Ala	Gly	Leu	Pro	Tyr	Met			
						260							265				270	
aag	cgc	ctg	aca	gtg	gcc	cac	cag	cta	cac	cac	agc	ggc	aag	tac	ggc	863		
Lys	Arg	Leu	Thr	Val	Ala	His	Gln	Leu	His	His	Ser	Gly	Lys	Tyr	Gly			
						275							280				285	
ggc	gcg	ccc	tgg	ggc	atg	ttc	ttg	ggc	cca	cag	gag	ctg	cag	cac	att	911		
Gly	Ala	Pro	Trp	Gly	Met	Phe	Leu	Gly	Pro	Gln	Glu	Leu	Gln	His	Ile			
						290							295				300	
cca	ggc	gcg	gcg	gag	gag	gtg	gag	cga	ctg	gtc	ctg	gaa	ctg	gac	tgg	959		
Pro	Gly	Ala	Ala	Glu	Glu	Val	Glu	Arg	Leu	Val	Leu	Glu	Leu	Asp	Trp			
						305							310				315	
tcc	aag	cgg	tag	ggcgccgaac	caggcacgct	gggttcacac	ctcatgcctg									1011		
Ser	Lys	Arg																
 320																		
tgataaggtg	tggttagagc	gatgcgtgtg	agacgggtat	gtcacggctg	actggtctga											1071		
tgaccaatgg	catcgcccat	gtctggctcat	cacgggctgg	ttgcctgggt	gaaggtgatg											1131		
cacatcatca	tgtgcggttg	gaggggctgg	cacagtgtgg	gctgaactgg	agcagttgtc											1191		
caggctggcg	ttgaatcagt	gaggggttgt	gattggcggg	tgtgaagcaa	tgactccgcc											1251		
catattctat	ttgtgggagc	tgagatgatg	gcatgcttgg	gatgtgcatg	gatcatggta											1311		
gtgcagcaaaa	ctatatccac	ctagggctgt	tggtaggatc	aggtgaggcc	ttgcacattg											1371		

catgatgtac tcgtcatggt gtgttggtga gaggatggat gtggatggat gtgtattctc 1431
agacgtagac cttgactgga ggcttgatcg agagagtggg ccgtattctt tgagagggga 1491
ggctcgtgcc agaaatggtg agtggatgac tgtgacgctg tacattgcag gcaggtgaga 1551
tgcactgtct cgattgtaaa atacattcag atgcaaaaaa aaaaaaaaaa aaaaaaa 1608

<210> 32

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 32

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly
1 5 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser
20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg
35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu
50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr
65 70 75 80

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu
85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg

76/358

100	105	110
Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly Val		
115	120	125
Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met		
130	135	140
Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu		
145	150	155
Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala		
165	170	175
His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys		
180	185	190
Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe		
195	200	205
Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe		
210	215	220
Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly		
225	230	235
Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val		
245	250	255
His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys		
260	265	270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly
 275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro
 290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser
 305 310 315 320

Lys Arg

<210> 33
 <211> 528
 <212> DNA
 <213> Erwinia uredovora

<220>
 <221> CDS
 <222> (1) .. (528)

<400> 33
 atg ttg tgg att tgg aat gcc ctg atc gtt ttc gtt acc gtg att ggc 48
 Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
 1 5 10 15
 atg gaa gtg att gct gca ctg gca cac aaa tac atc atg cac ggc tgg 96
 Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
 20 25 30
 ggt tgg gga tgg cat ctt tca cat cat gaa ccg cgt aaa ggt gcg ttt 144
 Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
 35 40 45
 gaa gtt aac gat ctt tat gcc gtg gtt ttt gct gca tta tcg atc ctg 192
 Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu

78/358

50		55		60		
ctg att tat ctg ggc agt aca gga atg tgg ccg ctc cag tgg att ggc	240					
Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly						
65		70		75		80
gca ggt atg acg gcg tat gga tta ctc tat ttt atg gtg cac gac ggg	288					
Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly						
		85		90		95
ctg gtg cat caa cgt tgg cca ttc cgc tat att cca cgc aag ggc tac	336					
Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr						
		100		105		110
ctc aaa cgg ttg tat atg gcg cac cgt atg cat cac gcc gtc agg ggc	384					
Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly						
		115		120		125
aaa gaa ggt tgt gtt tct ttt ggc ttc ctc tat gcg ccg ccc ctg tca	432					
Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser						
		130		135		140
aaa ctt cag gcg acg ctc cgg gaa aga cat ggc gct aga gcg ggc gct	480					
Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala						
		145		155		160
gcc aga gat gcg cag ggc ggg gag gat gag ccc gca tcc ggg aag taa	528					
Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys						
		165		170		175

<210> 34

<211> 175

<212> PRT

<213> Erwinia uredovora

<400> 34

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly
1 5 10 15

79/358

Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
20 25 30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe
35 40 45

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu
50 55 60

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly
65 70 75 80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly
85 90 95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr
100 105 110

Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
115 120 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser
130 135 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala
145 150 155 160

Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys
165 170 175

<210> 35

<211> 1520

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 35

ctcgagtacc gaggcggaac ggcaggaatg tttccctctc ttttagaggg caattcttta	60
tccaatgtca tgttgatgct agatatttct gtctcttata ataaggcgaa tacccatttt	120
tgaattgaag ttgagataaa aaaaaagggg gcccaatttg tcaacgcaa agagtcaagc	180
tttttctttg gctttagccg aacaatctaa gacttattgt ttttgaagat atttgacctt	240
ttctagatat tccttcaagt aaagcttttt tcgagttttt tttttttttc tttgtgaagg	300
atttattggt attggtatcc attttttatt ggaagacaag ataagttaat attgattttg	360
cttaaagatt aaaaggaaat cagaaaacga caataaaaaa tgtaacggac aaactatggt	420
gtcgattata agtctaaatc cttaaaaaat gacaacgagt tgctttcctc tgaaaacaat	480
tcttttgtct ttgcaagaaa ggtttctttt ttgtttgctt gcattactta aacatcaa	540
caaatgaaag gaataaagca gatttgaggg cgaataagga ttttctggtc aacaagatgt	600
gagtgcacc taaggaaact aatgccattc atttgtttta aaacgacatc aaagattgat	660
gatcaacagg attgagagag agaaaaagaa ctctgtgtcat ttatttctgt tgactgaaat	720
tttatattta gaaaaaatgt caaatctata gctttagcta tattacataa catttgaaat	780
aataataata aaaaaagaca cattagagac acttttcaaa ctctaaataa ctgtctataa	840
acacaaagaa aacaaagacc tctataacaa cttattagat ttttctcgta cttttgtcta	900
aagatgatgt attcttggtt tcccacactt ctttcatttg ttcttgatgc tactaaatat	960
acaaaatttc ttttttgcaa gagatattat tccaaaaatt ttcaaaaaga aatttttttc	1020
acaatagcag ttgatcgtgt aacccaaaga ggttctttgt tattttgcac ttccgctttg	1080

cggtgatgca tattcaaagt aatatatgga ataaacaacg tgtttaagca tgaaagaaag 1140
gaaacaaagg ccgctttgaa caaatgcata atatttcaga caaaatgat ctaaagcaag 1200
cagtaaatca aacaagaaac attgctgatt cgcgtagaa aacgataaaa gtctaataag 1260
ccactaagta tacttcaatg aactttttgt atgcttatgg tccaatcaga ccaataattt 1320
gtgaccattc ctgagggtggc tttgggtgatg cggaaacaga aaaaaatttt ctcaccaatc 1380
gatttaaaaa acaattttctg ctttgaacca aaactttttt tttctcttta atcattaact 1440
ttatcaagta tgtacctacc ctcaaagtcc tcaactcaagc acaattatgc taacattggt 1500
ccaccttctc tttagaatg 1520

<210> 36

<211> 16245

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 36

ccgggctggt tgcctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccc gcgttggtga tacctcgccg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gagggcgccg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgccgt	480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgctg gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaac cctcccggcc cgctaacgcg ggccctccat cccccaggg gctgcgcccc	660
tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgcca gccccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac cagggtgccg gcagtgaggg cggcggcctg	840
ggtggcggcc tgcccttcac ttccggcgtc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgccg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140
agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc	1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga	1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttccgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcccctg agttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100

aatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgcà tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattataattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cgggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggtg agtcggggca atcccgcaag 3360
gagggatgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcac 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcg cctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata ttctcgctatt 5100

ctggagcttg ttgtttatct cgggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtctctggggg ctatcttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcctggcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc ttgggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacgggt ctcccgtga cgcgctcccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctggt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900
gtcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgtga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggtt ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttgggt ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaate cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcggtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgtg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgcctttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct cgggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggtg tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt aatctataca 10800
atgctccata gactcacatt gatattgtcg aagatttcga tgctgactta gtagagcaac 10860
tacaaaagtt agcagagaag catgatttct taatctttga agaccgcaag tttgcagata 10920
tcggtatgtg aattctatct attttttttc tgatgtgtgc atggatgact catgatcata 10980
ttcttaggta atactgtcaa gcatcaatat ggcaagggcg tttacaagat tgcttcttgg 11040
tctcatatta ctaatgctca cacagttcct ggagaaggta ttatcaaggg acttgccgaa 11100
gtcggcctcc ctcttggtcg tggcttgctt ttgctagcag aaatgtcatc tcaaggtgca 11160
ttaactaagg gtatttacac tgccgaatct gtcaatatgg ctgccgcaa caaagatttc 11220
gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga ttttgttgtc 11280
atgtcgctg aagcttggcg taatcatggt catagctgtt tctgtgtga aattgttacc 11340

cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tgggggtgcct 11400
aatgagttag ctaactcaca ttaattgcgt tgcgctcact gcccgccttc cagtcgggaa 11460
acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagagggc gggttgcgta 11520
ttggggccaaa gacaaaaggg cgacattcaa ccgattgagg gaggggaaggt aaatattgac 11580
ggaaattatt cattaaaggt gaattatcac cgtcaccgac ttgagccatt tgggaattag 11640
agccagcaaa atcaccagta gcaccattac cattagcaag gccggaaacg tcaccaatga 11700
aaccatcgat agcagcaccg taatcagtag cgacagaatc aagtttgcct ttagcgtcag 11760
actgtagcgc gttttcatcg gcattttcgg tcatagcccc cttattagcg tttgccatct 11820
tttcataatc aaaatcaccg gaaccagagc caccaccgga accgcctccc tcagagccgc 11880
caccctcaga accgccaccc tcagagccac caccctcaga gccgccacca gaaccaccac 11940
cagagccgcc gccagcattg acaggaggcc cgatctagta acatagatga caccgcgcgc 12000
gataatttat cctagtttgc gcgctatatt ttgttttcta tcgctatta aatgtataat 12060
tgcggggactc taatcataaa aaccatctc ataaataacg tcatgcatta catgttaatt 12120
attacatgct taacgtaatt caacagaaat tatatgataa tcatcgcaag accggcaaca 12180
ggattcaatc ttaagaaact ttattgcaa atgtttgaac gatcggggat catccgggtc 12240
tgtggcgggg actccacgaa aatatccgaa cgcagcaaga tatcgcggtg catctcggtc 12300
ttgcctgggc agtcgccgcc gacgccgttg atgtggacgc cgggcccgat catattgtcg 12360
ctcaggatcg tggcgttgtg cttgtcggcc gttgctgtcg taatgatatc ggcaccttcg 12420
accgcctggt ccgcagagat cccgtggggc aagaactcca gcatgagatc cccgcgctgg 12480
aggatcatcc agccggcgct ccggaacacg attccgaagc ccaacctttc atagaaggcg 12540
gcggtggaat cgaaatctcg tgatggcagg ttgggcgtcg cttggtcggt catttcgaac 12600

cccagagtcc cgctcagaag aactcgtcaa gaaggcgata gaaggcgatg cgctgcgaat 12660
cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg ccaagctctt 12720
cagcaatata acgggtagcc aacgctatgt cctgatagcg gtccgccaca cccagccggc 12780
cacagtcgat gaatccagaa aagcggccat tttccaccat gatattcggc aagcaggcat 12840
cgccatgggt cacgacgaga tcatcgccgt cgggcatgcg cgccttgagc ctggcgaaca 12900
gttcggctgg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg acaagaccgg 12960
cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg aatgggcagg 13020
tagccggatc aagcgtatgc agccgccgca ttgcatcagc catgatggat actttctcgg 13080
caggagcaag gtgagatgac aggagatcct gccccggcac ttcgccaat agcagccagt 13140
cccttccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc gtcgtggcca 13200
gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcaccggac aggtcggctt 13260
tgacaaaaag aaccgggcgc ccctgcgctg acagccggaa cacggcggca tcagagcagc 13320
cgattgtctg ttgtgcccag tcatagccga atagcctctc cacccaagcg gccggagaac 13380
ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc agatccggtg cagattattt 13440
ggattgagag tgaatatgag actctaattg gataccgagg ggaatttatg gaacgtcagt 13500
ggagcatttt tgacaagaaa tatttgctag ctgatagtga ccttaggcga cttttgaacg 13560
cgcaataatg gtttctgacg tatgtgctta gtcattaaa ctccagaaac ccgcggctga 13620
gtggctcctt caacgttgcg gttctgtcag ttccaaacgt aaaacggctt gtcccgctc 13680
atcggcgggg gtcataacgt gactccctta attctccgct catgatcaga ttgtcgtttc 13740
ccgccttcag tttaaactat cagtgtttga caggatatat tggcgggtaa acctaagaga 13800
aaagagcgtt tattagaata atcggatatt taaaagggcg tgaaaagggt tatccgttcg 13860

tccatttgta tgtgcatgcc aaccacaggg ttccccagat ctggcgccgg ccagcgagac 13920
gagcaagatt ggccgcccgc cgaaacgata cgacagcgcg cccagcacag gtgcgcaggc 13980
aaattgcacc aacgcataca gcgccagcag aatgccatag tgggcgggtga cgtcgttcga 14040
gtgaaccaga tcgcgcagga ggcccggcag caccggcata atcaggccga tgccgacagc 14100
gtcgagcgcg acagtgtctca gaattacgat caggggtatg ttgggtttca cgtctggcct 14160
ccggaccagc ctccgctggt ccgattgaac gcgcggattc tttatcactg ataagttggt 14220
ggacatatta tgtttatcag tgataaagtg tcaagcatga caaagttgca gccgaataca 14280
gtgatccgtg ccgccctgga cctgttgaac gaggtcggcg tagacggtct gacgacacgc 14340
aaactggcgg aacggttggg ggttcagcag ccggcgcttt actggcactt caggaacaag 14400
cgggcgctgc tcgacgcact ggccgaagcc atgctggcgg agaatacatc gcattcgggtg 14460
ccgagagccg acgacgactg gcgctcattt ctgatcggga atgcccgcag cttcaggcag 14520
gcgctgctcg cctaccgcga tggcgcgcg atccatgccg gcacgcgacc gggcgccaccg 14580
cagatggaaa cggccgacgc gcagcttcgc ttctctcg aggcgggttt ttccggccggg 14640
gacgccgtca atgcgctgat gacaatcagc tacttctactg ttggggccgt gcttgaggag 14700
caggccggcg acagcgatgc cggcgagcgc ggccggcaccg ttgaacaggc tccgctctcg 14760
ccgctgttgc gggccgcgat agacgccttc gacgaagccg gtccggacgc agcgttcgag 14820
cagggactcg cggtgattgt cgatggattg gcgaaaagga ggctcgttgt caggaacgtt 14880
gaaggaccga gaaaggggtga cgattgatca ggaccgctgc cggagcgcaa ccactcact 14940
acagcagagc catgtagaca acatccccctc cccctttcca ccgcgtcaga cgcccgtagc 15000
agcccgttac gggctttttc atgccctgcc ctagcgtcca agcctcacgg ccgcgctcgg 15060
cctctctggc ggccttctgg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg 15120

tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 15180
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 15240
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 15300
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 15360
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 15420
tgtccgcctt tctcccttcg ggaagcgtgg cgctttttccg ctgcataacc ctgcttcggg 15480
gtcattatag cgattttttc ggtatatcca tccttttttcg cacgatatac aggattttgc 15540
caaagggttc gtgtagactt tccttggtgt atccaacggc gtcagccggg caggataggt 15600
gaagtaggcc caccgcgag cgggtgttcc ttcttctactg tcccttattc gcacctggcg 15660
gtgctcaacg ggaatcctgc tctgcgaggc tggccggcta ccgccggcgt aacagatgag 15720
ggcaagcgga tggctgatga aaccaagcca accaggaagg gcagcccacc tatcaagggtg 15780
tactgccttc cagacgaacg aagagcgatt gaggaaaagg cggcggcggc cggcatgagc 15840
ctgtcggcct acctgctggc cgtcggccag ggctacaaaa tcacgggcgt cgtggactat 15900
gagcacgtcc gcgagctggc ccgcatcaat ggcgacctgg gccgcctggg cggcctgctg 15960
aaactctggc tcaccgacga cccgcgcacg gcgcggttcg gtgatgccac gatcctcgcc 16020
ctgctggcga agatcgaaga gaagcaggac gagcttggca aggtcatgat gggcgtggtc 16080
cgcccgaggg cagagccatg acttttttag ccgctaaaac ggccgggggg tgccgctgat 16140
tgccaagcac gtcccatgc gtcctcatca gaagagcgac ttcgcggagc tgggtgaagta 16200
catcaccgac gagcaaggca agaccgagcg cctttgcgac gctca 16245

<210> 37

<211> 17877

<212> DNA

<213> Artificial

<220>

<223> Promotor

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 37

ccgggctggt tgcctcgcc gctgggctgg cggcgtcta tggcctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgccg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccg 480

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatatatt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc 660

tccggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcgga cggccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcgccc tgcccttcac ttccggccgc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgcggtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tctttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa ctattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatac cacgtcaaag 1560

ggtgacagca gggtcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttggggatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggtata taagctgggt ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcaactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc ttattgatc ttggggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcggggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcgggggca atcccgcaag 3360
gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaana accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgtcttg 3840
ccctgttcac caccgcgaac aagaaaatcc cgcgcgaggc gctgcaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttgaggc acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta caggaattc atatgggaga agtaccgcaa gctgtcgcgc acggccccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcatcgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcg cctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggctg tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttattt cgggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtcttggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgctggcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggcttttctca gccccagatc tggggtcgat cagccgggga tgcacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcgggttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta cctccgcga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgcgtcccg gactgatggg ctgcctgtat 6060
cgagtgggta ttttgtgccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cgggccacgc tggtttgccc cagcaggcga 6300
aaatcctgtt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcgggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttgggt ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaatata tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactatata gaataagata ggtggagagc ttataccgag ctcccaaata 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcggtga 7980

cctagctgat tctggagtga ccagaggggt catgacttga gcctaaaata cgccgctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaatctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttcgg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtccaa catttggtgc catattttcc tgcctctccc accagctgct 8760
cttttctttt ctctttcttt tccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttatcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatgggtt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gcccattecg 9360
accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agtccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatttg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttectac taagatggtg tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800
tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860
aagataagtt aatattgatt ttgcttaaag attaaaagga aatcagaaaa cgacaataaa 10920
aaatgtaacg gacaaactat ggtgtcgtt ataagtctaa atccttaaaa aatgacaacg 10980
agttgctttc ctctgaaaac aattcttttg tctttgcaag aaagggttct tttttgtttg 11040
cttgcattac ttaaacaatca aatcaaatga aaggaataaa gcagatttga gggcgaataa 11100
ggattttctg gtcaacaaga tgtgagtgac acctaggaa cttaatgcca ttcatttggt 11160
ttaaaccgac atcaaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220
catttatctc tgttgactga aattttatat ttagaaaaaa tgtcaaatct atagctttag 11280
ctatattaca taacatttga aataataata ataaaaaag acacattaga gacacttttc 11340
aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400
gatttttctc gtacttttgt ctaaagatga tgtattcttg ttatcccaca cttctttcat 11460
ttgttcttga tgctactaaa tatacaaaat ttcttttttg caagagatat tattccaaaa 11520
attttcaaaa agaaattttt ttcacaatag cagttgatcg tgtaacccaa agaggttctt 11580
tgttattttg cacttccgct ttgcggtgat gcatattcaa agtaatatat ggaataaaca 11640
acgtgtttta gcatgaaaga aaggaaacaa aggccgcttt gaacaaatgc ataatatctc 11700
agacaaaaat gatctaaagc aagcagtaaa tcaaacaaga aacattgctg attcgcgtta 11760
gaaaacgata aaagtctaata aagccactaa gtataacttca atgaactttt tgtatgctta 11820

tgggtccaatc agaccaataa tttgtgacca ttcttgaggt ggctttggtg atgcggaaac 11880
agaaaaaat tttctcacca atcgatttaa aaaacaattt ctgctttgaa ccaaaacttt 11940
ttttttctct ttaatcatta actttatcaa gtatgtacct accctcaaag tcttcaactca 12000
agcacaatta tgctaacatt gttccacctt ctcttttagaa atgctgtcga agctgcagtc 12060
aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac atcacgcggc ccaaagtctg 12120
cctgcatgct cagcgggtgct cgtagttcg gctgcgagtg gcagcaccac agacagagga 12180
ggcgctggga accgtgcagg ctgccggcgc gggcgatgag cacagcgccg atgtagcact 12240
ccagcagctt gaccgggcta tcgcagagcg tcgtgcccgg cgcaaacggg agcagctgtc 12300
ataccaggct gccgccattg cagcatcaat tggcgtgtca ggcatcgcca tcttcgccac 12360
ctacctgaga tttgccatgc acatgaccgt gggcggcgca gtgccatggg gtgaagtggc 12420
tggcactctc ctcttggtgg ttggtggcgc gctcggcatg gagatgtatg cccgctatgc 12480
acacaaagcc atctggcatg agtcgcctct gggctggctg ctgcacaaga gccaccacac 12540
acctgcgact ggacccttg aagccaacga cttgtttgca atcatcaatg gactgcccgc 12600
catgctcctg tgtaccttg gcttctggct gcccaacgtc ctggggggcg cctgctttgg 12660
agcggggctg ggcatacgc tatacggcat ggcataatg tttgtacacg atggcctggt 12720
gcacaggcgc tttcccaccg ggcccatcgc tggcctgcc tacatgaagc gcctgacagt 12780
ggcccaccag ctacaccaca gcggcaagta cgggtggcgc cctgggggta tgttcttggg 12840
tccacaggag ctgcagcaca ttccagggtc ggcggaggag gtggagcgac tggctcctgga 12900
actggactgg tccaagcggc agaagcttgg cgtaatcatg gtcatagctg tttcctgtgt 12960
gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 13020
cctgggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 13080

tccagtcggg aaacctgtcg tgccagctgc attaataaat cggccaacgc gcgggggagag 13140
gcggtttgcg tattggggcca aagacaaaag ggcgacattc aaccgattga gggaggggaag 13200
gtaaatattg acggaaatta ttcattaaag gtgaattatc accgtcaccg acttgagcca 13260
tttgggaatt agagccagca aaatcaccag tagcaccatt accattagca aggccggaaa 13320
cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc 13380
ctttagcgtc agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag 13440
cgtttgccat cttttcataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc 13500
cctcagagcc gccaccctca gaaccgccac cctcagagcc accaccctca gagccgccac 13560
cagaaccacc accagagccg ccgccagcat tgacaggagg cccgatctag taacatagat 13620
gacaccgcgc gcgataattt atcctagttt gcgcgctata ttttgtttcc tatcgcgat 13680
taaagtata attgcgggac tctaataata aaaaccatc tcataaataa cgatcatgat 13740
tacatgttaa ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca 13800
agaccggcaa caggattcaa tcttaagaaa ctttattgcc aaatgtttga acgatcgggg 13860
atcatccggg tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcgcg 13920
tgcatctcgg tcttgccctg gcagtcgccg ccgacgccgt tgatgtggac gccggggccc 13980
atcatattgt cgctcaggat cgtggcggtg tgcttgctcg ccgttgctgt cgtaatgata 14040
tcggcacctt cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga 14100
tccccgcgct ggaggatcat ccagccggcg tcccggaaaa cgattccgaa gcccaacctt 14160
tcatagaagg cggcggtgga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg 14220
gtcatttcga accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga 14280
tgcgctgcga atcgggagcg gcgataccgt aaagcacgag gaagcgggtca gcccatcgc 14340

cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtcctgatag cgggtccgcca 14400
caccagccg gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg 14460
gcaagcaggc atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cggccttga 14520
gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcctcctgat 14580
cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat gcgatgtttc gcttggtggg 14640
cgaatgggca ggtagccgga tcaagcgtat gcagccgccc cattgcatca gccatgatgg 14700
atactttctc ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgccc 14760
atagcagcca gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc 14820
ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg 14880
acaggtcggt cttgacaaaa agaaccgggc gcccttgccg tgacagccgg aacacggcgg 14940
catcagagca gccgattgtc tgttggtgcc agtcatagcc gaatagcctc tccaccaag 15000
cggccggaga acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg 15060
tgcagattat ttggattgag agtgaatatg agactctaata tggataccga ggggaattta 15120
tggaacgtca gtggagcatt ttgacaaga aatatttgct agctgatagt gaccttaggc 15180
gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 15240
acccgagggt gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc 15300
ttgtcccgcg tcctcggcgg gggtcataac gtgactccct taattctccg ctcatgatca 15360
gattgtcggt tccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt 15420
aaacctaaga gaaaagagcg tttattagaa taatcgata tttaaaaggg cgtgaaaagg 15480
tttatccgtt cgtccatttg tatgtgcatg ccaaccacag ggttccccag atctggcgcc 15540
ggccagcgag acgagcaaga ttggccgccc cccgaaacga tccgacagcg cgcccagcac 15600

aggtgcgag gcaaattgca ccaacgcata cagcgccagc agaatgccat agtgggcggt 15660
gacgtcgttc gagtgaacca gatcgcgagc gaggcccggc agcaccggca taatcaggcc 15720
gatgccgaca gcgtcgagcg cgacagtgtc cagaattacg atcaggggta tgttgggttt 15780
cacgtctggc ctccggacca gcctccgctg gtccgattga acgcgcggat tctttatcac 15840
tgataagttg gtggacatat tatgtttatc agtgataaag tgtcaagcat gacaaagttg 15900
cagccgaata cagtgatccg tgccgccctg gacctgttga acgaggtcgg cgtagacggt 15960
ctgacgacac gcaaactggc ggaacggttg ggggttcagc agccggcgct ttactggcac 16020
ttcaggaaca agcgggcgct gctcgacga ctggccgaag ccatgctggc ggagaatcat 16080
acgcattcgg tgccgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgcccgc 16140
agcttcaggc aggcgctgct cgcctaccgc gatggcgcg ccatccatgc cggcacgcga 16200
ccgggcgcac cgcagatgga aacggccgac gcgcagcttc gcttcctctg cgaggcgggt 16260
ttttcggccg gggacgccgt caatgcgctg atgacaatca gctacttcac tgttggggcc 16320
gtgcttgagg agcaggccgg cgacagcgat gccggcgagc gcggcggcac cgttgaacag 16380
gctccgctct cgccgctgtt gcgggcccgc atagacgcct tcgacgaagc cggctccggac 16440
gcagcgttcg agcagggact cgcggtgatt gtcgatggat tggcgaaaag gaggctcggt 16500
gtcaggaacg ttgaaggacc gagaaagggt gacgattgat caggaccgct gccggagcgc 16560
aaccactca ctacagcaga gccatgtaga caacatcccc tcccccttcc caccgcgtca 16620
gacgcccgtg gcagcccgt acgggctttt tcatgccctg ccctagcgtc caagcctcac 16680
ggccgcgctc ggccctctctg gcggccttct ggcgctcttc cgcttcctcg ctactgact 16740
cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaag gcggtaatac 16800

ggttatccac agaatcagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 16860
aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 16920
acgagcatca caaaaatcga cgctcaagtc agagggtggcg aaacccgaca ggactataaa 16980
gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 17040
ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgcttttc cgctgcataa 17100
ccctgcttcg gggtcattat agcgattttt tcggtatata catccttttt cgcacgatat 17160
acaggatttt gccaaaggg tctgttagac tttccttggg gtatccaacg gcgtcagccg 17220
ggcaggatag gtgaagtagg cccacccgcg agcgggtggt ccttcttcac tgtcccttat 17280
tcgcacctgg cgggtgctcaa cgggaatcct gctctgcgag gctggccggc tacgccggc 17340
gtaacagatg agggcaagcg gatggctgat gaaaccaagc caaccaggaa gggcagccca 17400
cctatcaagg tgtactgcct tccagacgaa cgaagagcga ttgaggaaaa ggcggcgggc 17460
gccggcatga gcctgtcggc ctacctgctg gccgtcggcc agggctacaa aatcacgggc 17520
gtcgtggact atgagcacgt ccgcgagctg gccgcacatca atggcgacct gggccgcctg 17580
ggcggcctgc tgaaactctg gctcaccgac gaccgcgcga cggcgcggtt cggatgatgcc 17640
acgatcctcg ccctgctggc gaagatcgaa gagaagcagg acgagcttgg caaggatcatg 17700
atgggcgtgg tccgcccag ggcagagcca tgactttttt agccgctaaa acggccgggg 17760
ggcgcgcgtg attgccaagc acgtcccat gcgctccatc aagaagagcg acttcgcgga 17820
gctggatgaag tacatcaccg acgagcaagg caagaccgag cgcctttgcg acgctca 17877

<210> 38

<211> 17238

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 38

ccgggctggt tgcctcgcc gctgggctgg cggcgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgccgt	480
ttttcgcca ccgctaacct gtcttttaac ctgcttttaa accaatatit ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc	660

tgggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgcgga cgcccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcggcc tgcccttcac ttccggcgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttataa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatacttata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gaggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcccttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgct ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctgggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atccccgaag 3360
gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gacgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac cgcgcgaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccgggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta caggaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgagge agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcg cctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagccccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatth cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatttgcg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gccccagatc tggggctgat cagccgggga tgcatcaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgctga cgccgtcccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cggccacgc tggtttgccc cagcaggcga 6300

aatcctgtt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggctcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaaggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcgggtacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtgggtg caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttgggt ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440
tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaa gggcgttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaata 7920
tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcggtga 7980
cctagctgat tctggagtga cccagaggggt catgacttga gcctaaaata cgcgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcattcatcg 8280

agaatatgga gcttcatcga atcacccgca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtccttg gtaggcagct ttgccccgtc tgtccgccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttggtgc catattttcc tgcctctccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgtctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcctgcagc ttgactaaca gctacccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcaagaat ctgctgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcctcggc gcgctccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcagc cggctcggga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggg 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcttgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgagggc atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc a cactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgac agtccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgtccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttgttt agaggtaatc cttcttttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ctaccgcttg 10800
gaccagtcca gttccaggac cagtcgctcc acctcctccg ccgcacctgg aatgtgctgc 10860
agctcctgtg gacccaagaa cataccccag ggcgcgccac cgtacttgcc gctgtggtgt 10920
agctggtggg ccactgtcag gcgcttcattg tagggcaggc cagcgatggg cccggtggga 10980
aagcgctgt gcaccaggcc atcgtgtaca aacatatatg ccatgcccga tagcgtgatg 11040
cccagccccg ctccaaagca ggccgcccc aggacgttg gcagccagaa gccaaaggta 11100
cacaggagca tggcgggcag tccattgatg attgcaaaca agtcgttggc ttcaaagggt 11160
ccagtgcgag gtgtgtggtg gctcttgctg agcagccagc ccagaggcga ctcatgccag 11220
atggctttgt gtgcatagcg ggcatacatc tccatgccga gcgcgccacc aaccaccaag 11280
aggagagtgc cagccacttc accccatggc actgcgcgc ccacggtcat gtgcatggca 11340
aatctcaggt aggtggcgaa gatggcaatg cctgacacgc caattgatgc tgcaatggcg 11400
gcagcctggt atgacagctg ctcccgtttg cgccgggcac gacgctctgc gatagcccgg 11460
tcaagctgct ggagtgtac atcggcgtg tgetcatcgc ccgcgccggc agcctgcacg 11520
gttcccagcg cctcctctgt ctgtggtgct gccactcgca gccgaactaa cgagcaccgc 11580
tgagcatgca ggcagacttt gggccgcgtg atgtcgcggg ctagttcaac gcggcggggc 11640
ttgacgctga ttgactgcag cttcgacagc atagagataa aataaaaaga gaagaaaaga 11700
aagtttgtac aatttctttt tgtttatata acatacacgc tatgtcaaca tttagaataa 11760
gggggaaaaa atcttccatc atattcgaat gcacaagatt atttctttgt tcgctctttt 11820
tggtcggggtc atcgagattt agagtgtaat caaagatact gtcactctga gagcgttgca 11880
caggctgctg tttgccaaat tggatgtttg ccgaattagt aaaatacgca agcatttctt 11940

acctttccgc tcccttttcc taattctccc aaagactaaa tgaggaaaga taaaggacaa 12000
agaaaatgta aagacaaaga aattgaaaac gatataaact tgcagcacgt aagaccaaag 12060
caaattggta actattcttg tgtacaaaca tgtataaaaa aaaacttttt tttgctcctg 12120
gaggacaaaa tttcaaactc cttgaagaag attgcttgta tatctatcat atgcatatat 12180
catatcgatg gaaaaagaaa gtcaggcatg tatttataaa aagaagaatg tgccatgctt 12240
ccgaatttct tttcactttc ttttccttat ctattttaat ctcaagcttg gcgtaatcat 12300
ggtcatagct gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag 12360
ccggaagcat aaagtgtaaa gcctgggggtg cctaagagt gagctaactc acattaattg 12420
cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12480
tcggccaacg cgcggggaga ggcggtttgc gtattgggcc aaagacaaaa gggcgacatt 12540
caaccgattg agggagggaa ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12600
caccgtcacc gacttgagcc atttgggaaat tagagccagc aaaatcacca gtagcaccat 12660
taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12720
tagcgacaga atcaagtttg ccttttagcgt cagactgtag cgcgttttca tcggcatttt 12780
cggtcatagc ccccttatta gcgtttgcca tcttttcata atcaaaatca ccggaaccag 12840
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcca ccctcagagc 12900
caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12960
gcccgatcta gtaacataga tgacaccgcg cgcgataatt taccctagtt tgcgcgctat 13020
attttgtttt ctatcgcgta ttaaattgat aattgcggga ctctaatacat aaaaacccat 13080
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13140

aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13200
caaagtgttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13260
gaacgcagca agatatcgcg gtgcatctcg gtcttgcccg ggcagtcgcc gccgacgccg 13320
ttgatgtgga cgccggggccc gatcatattg tcgctcagga tcgtggcggt gtgcttgtcg 13380
gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccgtgg 13440
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaaa 13500
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13560
aggttgggcg tcgcttggtc ggtcatttcg aacccagag tcccgcctcag aagaactcgt 13620
caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13680
ggaagcggtc agcccattcg ccgccaagct cttcagcaat atcacgggta gccaacgcta 13740
tgtcctgata gcggtccgcc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13800
cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13860
cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13920
cttcgtccag atcactctga tcgacaagac cggcttccat ccgagtacgt gctcgtcga 13980
tgcgatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 14040
gcattgcac agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 14100
cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14160
gcacagctgc gcaaggaacg ccgctcgtgg ccagccacga tagccgcgct gcctcgtcct 14220
gcagttcatt cagggcaccg gacaggtcgg tcttgacaaa aagaaccggg cgcccctgcg 14280
ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14340
cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14400

tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatattgc 14520
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg cttgtcccgc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gctcatgatc agattgtcgt ttccgcctt cagtttaaac tatcagtgtt 14760
tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataatcggat 14820
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaaccaca 14880
gggttcccca gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgcccagca cagggtgcga ggcaaattgc accaacgcat acagcgccag 15000
cagaatgcca tagtggcggt tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 15060
cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120
gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15240
gtgtcaagca tgacaaagt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15300
aacgaggtcg gcgtagacgg tctgacgaca cgaaaactgg cggaacggtt gggggttcag 15360
cagccggcgc ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15420
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15480
tttctgatcg ggaatgccc cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15540
cgcatccatg ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt 15600
cgcttcctct gcgaggcggg tttttcgcc ggggacgccg tcaatgcgct gatgacaatc 15660

agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15720
cgcgggcgga ccgttgaaca ggctccgctc tcgccgctgt tgcgggccgc gatagacgcc 15780
ttcgacgaag ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgtcgatgga 15840
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15900
tcaggaccgc tgccggagcg caaccctc actacagcag agccatgtag acaacatccc 15960
ctcccccttt ccaccgcgtc agacgcccg agcagccgc tacgggcttt ttcattgccct 16020
gccctagcgt ccaagcctca cggccgcgt cggcctctct ggcggccttc tggcgctctt 16080
ccgcttcctc gctcactgac tcgctgcgt cggtcgttcg gctgcggcga gcggtatcag 16140
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16200
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgctttt 16260
tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc 16320
gaaacccgac aggactataa agataaccagg cgtttcccc tggaagctcc ctctgctcgt 16380
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 16440
tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16500
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16620
tccttcttca ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga 16680
ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16800
attgaggaaa aggcggcgcc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16860
cagggtaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccccatc 16920

aatggcgacc tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc 16980
acggcgcggt tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggtcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 17100
tagccgctaa aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat 17160
caagaagagc gacttcgagg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17220
gcgcctttgc gacgctca 17238

<210> 39

<211> 17238

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 39

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgccg 120

aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggccca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc 660

tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgca cgcccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcgggc tgccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatata gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcaccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta tttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cgggtctaagg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgagggc gctgcaaaac aagggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggg gtggcagcag gtgttgaggc acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccgggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcgggtgtcg tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggctcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgcgg acggccccgac 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggg agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcattgcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gtcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccca 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcggggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctgggtga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatth cggctctaccg cctgccgggc ggggtcgcg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatttgcg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcctggcaa ctggcggccg gaggacttct gtcggtcca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttgggtccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcaact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgct tacaacggct ctcccgtga cgccgtccc gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatcc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920
tgtccagatc atgggtgacc ggtgcctgga tcttcctata gaatcatcct tattegttga 7980
cctagctgat tctggagtga cccagagggt catgacttga gcctaaaadc cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccc gtgtgtcggc ggggttgaca 8700
aggctgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctaccccg 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcggtgcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatgggtt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacagg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gcccatcgg 9360
accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct cgggcgaaga gaagaatagc ttagcagagc tattttcatt ttcggggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acaattgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc cactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800
taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860
tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920
ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca aagatactgt 10980
catctcgaga gcgttgaca ggctgctgtt tgccaaattg gatgtttgcc gaattagtaa 11040
aatacgcaag catttcttac ctttccgctc ctttttcta attctccaa agactaaatg 11100
aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160
cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaacatg tataaaaaaa 11220
aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgata 11280
tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta tttataaaaa 11340

gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct attttaatct 11400
catgctgtcg aagctgcagt caatcagcgt caaggcccg cgcgttgaac tagcccgga 11460
catcacgcgg cccaaagtct gcctgcatgc tcagcggcgc tcgttagttc ggctgcgagt 11520
ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgcggcg cgggcgatga 11580
gcacagcgcc gatgtagcac tccagcagct tgaccgggct atcgagagc gtcgtgccc 11640
gcgcaaacgg gagcagctgt cataccaggc tgccgccatt gcagcatcaa ttggcgtgtc 11700
aggcattgcc atcttcgcca cctacctgag atttgccatg cacatgaccg tgggcggcg 11760
agtgccatgg ggtgaagtgg ctggcactct cctcttgggtg gttggtggcg cgctcgcat 11820
ggagatgtat gcccgctatg cacacaaagc catctggcat gagtcgcctc tgggctggct 11880
gctgcacaag agccaccaca cacctgcac tggacccttt gaagccaacg acttgtttgc 11940
aatcatcaat ggactgccc ccatgctcct gtgtaccttt ggcttctggc tgcccaacgt 12000
cctggggggcg gcctgcttg gagcggggct gggcatcacg ctatacggca tggcatatat 12060
gtttgtacac gatggcctgg tgcacaggcg ctttcccacc gggcccatcg ctggcctgcc 12120
ctacatgaag cgctgacag tggcccacca gctacaccac agcggcaagt acggtggcg 12180
gccctgggggt atgttcttgg gtccacagga gctgcagcac attccaggtg cggcggagga 12240
ggtggagcga ctggtcctgg aactggactg gtccaagcg tagaagcttg gcgtaatcat 12300
ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag 12360
ccggaagcat aaagtgtaaa gcctgggggtg cctaagagt gagctaactc acattaattg 12420
cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12480
tcggccaacg cgcggggaga ggcgggttgc gtattgggccc aaagacaaaa gggcgacatt 12540
caaccgattg agggagggaa ggtaaatatt gacggaaatt attcattaaa ggtgaattat 12600

caccgtcacc gacttgagcc atttgggaat tagagccagc aaaatcacca gtagcaccat 12660
taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12720
tagcgacaga atcaagtttg ccttttagcgt cagactgtag cgcgttttca tcggcatttt 12780
cggtcatagc ccccttatta gcgtttgcca tcttttcata atcaaaatca ccggaaccag 12840
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcca ccctcagagc 12900
caccaccctc agagccgcca ccagaaccac caccagagcc gccgccagca ttgacaggag 12960
gcccgatcta gtaacataga tgacaccgcg cgcgataatt taccctagtt tgcgcgctat 13020
attttgtttt ctatcgcgta ttaaattgat aattgcggga ctctaatacat aaaaaccat 13080
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13140
aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13200
caaattgttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13260
gaacgcagca agatatcgcg gtgcatctcg gtcttgccctg ggcagtcgcc gccgacgccg 13320
ttgatgtgga cgccggggcc gatcatattg tcgctcagga tcgtggcggt gtgcttgctg 13380
gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccgtag 13440
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccgga 13500
acgattccga agcccaacct ttcatagaag gcggcggttg aatcgaaatc tcgtgatggc 13560
aggttgggcg tcgcttggtc ggtcatttcg aaccccagag tccgctcag aagaactcgt 13620
caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13680
ggaagcggtc agcccatcgc ccgccaagct cttcagcaat atcacgggta gccaacgcta 13740
tgtcctgata gcgggtccgc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13800
cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13860

cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13920
cttcgtccag atcctcctga tcgacaagac cggcttccat ccgagtacgt gctcgtcga 13980
tgcgatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 14040
gcattgcac agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 14100
cctgccccgg cacttcgccc aatagcagcc agtccttcc cgcttcagtg acaacgtcga 14160
gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgt gcctcgtcct 14220
gcagttcatt cagggcaccg gacaggtcgg tcttgacaaa aagaaccggg cgccccctgcg 14280
ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14340
cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14400
tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460
ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatttgc 14520
tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14580
ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640
cagttccaaa cgtaaaacgg cttgtccgc gtcacggcg ggggtcataa cgtgactccc 14700
ttaattctcc gtcatgatc agattgtcgt tccccgctt cagtttaaac tatcagtgtt 14760
tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataatcggat 14820
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gcccaaccaca 14880
gggttcccca gatctggcg cggccagcga gacgagcaag attggccgcc gcccgaaacg 14940
atccgacagc gcgccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 15000
cagaatgcca tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccg 15060
cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120

gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15240
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15300
aacgaggctcg gcgtagacgg tctgacgaca cgcaaactgg cggaacgggt gggggttcag 15360
cagccggcgc tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15420
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15480
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15540
cgcattccatg ccggcacgcg accggggcgca ccgcagatgg aaacggccga cgcgcagctt 15600
cgcttctctt gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc 15660
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15720
cgcggcggca ccgttgaaca ggctccgctc tcgccgctgt tgccggccgc gatagacgcc 15780
ttcgacgaag ccggtccgga cgcagcgttc gagcagggac tcgcggtgat tgctgatgga 15840
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15900
tcaggaccgc tgccggagcg caaccactc actacagcag agccatgtag acaacatccc 15960
ctcccccttt ccaccgcgtc agacgcccggt agcagccgcg tacgggcttt ttcattgcct 16020
gccctagcgt ccaagcctca cggccgcgct cggcctctct ggccggccttc tggcgctctt 16080
ccgcttcttc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag 16140
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16200
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcggtt ctggcgctttt 16260
tccataggct ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagaggtggc 16320
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctctgctgct 16380

ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 16440
tggcgctttt ccgctgcata accctgcttc ggggtcatta tagcgatttt ttcggtatat 16500
ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16560
tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16620
tcctttcttca ctgtccctta ttcgcacctg gcggtgctca acgggaatcc tgctctgcga 16680
ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16740
ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16800
attgaggaaa aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16860
cagggctaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcac 16920
aatggcgacc tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgaccgcgc 16980
acggcgcggt tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag 17040
gacgagcttg gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 17100
tagccgctaa aacggccggg ggggtgcgct gattgccaag cacgtcccca tgcgctccat 17160
caagaagagc gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17220
gcgcctttgc gacgctca 17238

<210> 40

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 40

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60

cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240

tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agcttttctca 300

gccagggcca gcccaagacc gacaaggcct cctccagaa cgccgagaag aactggaggg 360

gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 420

atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

tttggatttg ctttccaggc tgagactcta gcttgaggca tagagggtcc tttggctttc 540

aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600

gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660

ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720

tccgatagct ctgcaaaggc cgttgcacia tgcaagtcaa acgttgctag cagttccagg 780

tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacaaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgacct agagggtcat gacttgagcc taaaatccgc cgcctccacc attttagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttcgac tcggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt accggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctctg ctttgcccgg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgtg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctgt ccgcccgtg tgtcggcggg gttgacaagg tcgttgcgtc 1920

agccaacat ttgttgccat attttctgc tctcccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatcaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctccgagggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcgctg gatatgtcct gcgggtaaata agctgcgccg atggtttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcac tcccgccgtg cacaggggtgt cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
ggatcaatata ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga tttcggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
ttcgggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgacag atcgccgcgg 2940
ctccgggcgt atatgctccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctcc atccggagcc 3060
gggactgtcg ggcgtaacaa aatcgcccg cgaagcgcgg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctcgctccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagtg 3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggtcc acgcgactat 3600
atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tcottgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780
gatggtatac aatagtaacc atgcatgggt gcctagtga tgctccgtaa cacccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacagggtaca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020
agaaagtga aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080
ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140
ggagtttgaa attttgtcct ccaggagcaa aaaaaagttt ttttttatac atgtttgtac 4200
acaagaatag ttaccaattt gctttggtct tacgtgctgc aagtttatat cgttttcaat 4260
ttctttgtct ttacattttc ttgtccttt atctttcctc atttagtctt tgggagaatt 4320
aggaaaagg agcggaagg taagaaatgc ttgcgtattt tactaattcg gcaaacatcc 4380
aatttggaac acagcagcct gtgcaacgct ctcgagatga cagtatctt gattacactc 4440
taaactctga tgaccgacc aaaagagcg acaaagaaa taatcttgtg cattcgaata 4500
tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620

gctgcagtca atcagcgtca aggcccgccg cggtgaacta gcccgcgaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcgggtgctc gttagtctcg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaacggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgagc tgccatgggg 4980

tgaagtggct ggcactctcc tcttggtggt tggtagcgcg ctccggcatgg agatgtatgc 5040

ccgctatgca cacaagcca tctggcatga gtcgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctcgcactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgcccgc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tgggggcggc 5220

ctgctttgga gcggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280

tggcctgggt cacaggcgt ttcccaccgg gcccatcgct ggcctgccct acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccagggtcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcggta gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctaggtgatc aaagactggt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaaga ttgctctgtt 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttggtg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtcttttaa tgtaattta gtagagtgtg 6000

aagcctgcat tttttttgta cgcataaaca atgaattcac cccgcttctg gtttttaa 6060

aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttggtgt 6120

ggagtcacag gcttgtcttc agattgatag aggttggtata cactcaacag agcaatcttg 6180

gcacgttcgc ttcttttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttggtcaa gcagctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcggtgcc ttgctgtatt 6420

tgttctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac aaaaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600

tattcaatcc ataaatgaat tattttttgga caaaaaagaa agattatgcc tgattttctc 6660

tatttttttt ttttttacia ctccaccaat actttctagc ccagcttggc gtaatcatgg 6720

tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc 6780

ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg 6840

ttgcgctcac tgcccgttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900

ggccaacgcg cggggagagg cggtttgcgt attgggcca agacaaaagg gcgacattca 6960

accgattgag ggaggggaagg taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080
ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140
gcgacagaat caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg 7200
gtcatagccc ccttattagc gtttgccatc ttttcataat caaatcacc ggaaccagag 7260
ccaccaccgg aaccgcctcc ctacagagccg ccaccctcag aaccgccacc ctacagagcca 7320
ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc 7380
ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat 7440
tttgttttct atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500
cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560
ttatatgata atcatcgcaa gaccggcaac aggattcaat cttaagaaac ttatttgcca 7620
aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680
acgcagcaag atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt 7740
gatgtggacg cggggcccgga tcatattgtc gtcaggatc gtggcgttgt gcttgtcggc 7800
cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tccgtgggc 7860
gaagaactcc agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920
gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980
gttgggcgtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040
agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100
aagcggtcag cccattcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg 8160
tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220
ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280

tcgggcatgc gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340
tcgtccagat catcctgata gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400
cgatgtttcg cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460
attgcatcag ccatgatgga tacttttctcg gcaggagcaa ggtgagatga caggagatcc 8520
tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgc aacgtcgagc 8580
acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc 8640
agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8700
gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760
aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820
cgaaacgata cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880
ggataccgag ggggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940
gctgatagtg accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000
agctcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060
gttccaaacg taaaacggct tgtccgcgt catcggcggg ggtcataacg tgactccctt 9120
aattctccgc tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 9180
acaggatata ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat 9240
ttaaagggc gtgaaaaggc ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300
gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgccgc ccgaaacgat 9360
ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgcagca 9420
gaatgccata gtgggcgggtg acgtcgttcg agtgaaccag atcgcgcagg aggccgggca 9480
gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga 9540

tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600

cgcgcggtatt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720

cgaggctcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780

gccggcgctt tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc 9840

catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt 9900

tctgatcggg aatgcccga gcttcaggca ggcgctgctc gcctaccgag atggcgcgcg 9960

catccatgcc ggcacgcgac cgggcgccacc gcagatggaa acggccgacg cgcagcttcg 10020

cttcctctgc gaggcgggtt ttccggccgg ggacgccgctc aatgcgctga tgacaatcag 10080

ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140

cggcgccacc gttgaacagg ctccgctctc gccgctgttg cgggcccga tagacgcctt 10200

cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260

ggcgaaaagg aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc 10320

aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380

ccccctttcc accgcgtcag acgcccgtag cagcccgtta cgggcttttt catgccctgc 10440

cctagcgtcc aagcctcac gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500

gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgccggcgagc ggtatcagct 10560

cactcaaagg cggtaatagc gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680

cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aaccgcagag gactataaag ataccaggcg ttccccctg gaagctccct cgtgcgctct 10800

cctgttccga cctgcccgt taccggatac ctgtccgcct ttctcccttc ggggaagcgtg 10860
gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc 10920
atcctttttc gcacgatata caggattttg ccaaagggtt cgtgtagact ttccttggtg 10980
tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc 11040
cttctttcact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100
ctggccgggt accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160
aaccaggaag ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220
tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280
gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340
tggcgacctg ggccgcctgg gcggcctgct gaaactcttg ctcaccgacg acccgcgcac 11400
ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460
cgagcttggc aaggtcatga tgggcgtggt ccgcccagg gcagagccat gactttttta 11520
gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580
agaagagcga cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc 11640
gcctttgcga cgctcaccgg gctggttgcc ctgcgccgtg ggctggcggc cgtctatggc 11700
cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760
tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820
tgagggggccg actcaccgg cgcggcgttg acagatgagg ggcaggctcg atttcggccg 11880
gcgacgtgga gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc 11940
ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 12000
gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060

tgaggggagc acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 12120

aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180

atatttataa accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg 12240

aaggggggtg ccccccttc tcgaaccctc ccggcccgtt aacgcggggc tcccatcccc 12300

ccaggggctg cggccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt 12360

ccttgccatt gccgggatcg gggcagtaac gggatgggag atcagcccga gcgcgacgcc 12420

cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480

tgaggggggc ggctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540

cttcattggc gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600

cgtgctcgtg ttcgggggtg cgataaaccc agcgaaccat ttgaggtgat aggtgaagatt 12660

ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720

ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat 12780

attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840

tttcaggggg caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaagca 12900

taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960

ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc 13020

gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080

tgccaggtgc tgcctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140

gattacgtgc agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca 13200

tatcaccagc tcaaaggggtg acagcaggct cataagacgc ccagcgtcg ccatagtgcg 13260

ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca 13320

gcgctggcgc gatttagccc cgacatagcc ccactgttcg tccatttccg cgagacgat 13380
gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggctgagtt ttttaagtga 13440
cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca 13500
ttcatggcca tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac 13560
tgcagttgcc atgttttacg gcagtgcgag cagagatagc gctgatgtcc ggcgggtgctt 13620
ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680
gccactggag cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc 13740
cataattgtg gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac 13800
aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860
gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920
ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980
accgctgcgt aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040
aaaatgaaaa cctatattta aaaatgacgg acagccggtg taaagggacc acctatgatg 14100
tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc 14160
tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc 14220
tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280
agtgcacag gctctttcac tccatcgaca tatcggttg tccctatacg aatagcttag 14340
acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg 14400
aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460
cggaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520

ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580
agtggatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700
atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760
caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820
aagtatttgg gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac 14880
gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940
gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc 15000
ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060
gaactgatcg acgcggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120
atgcgtgcgc cccgcgaaac cttccagtcc gtcggctcga tgggccagca agctacggcc 15180
aagatcgagc ggcacagcgt gcaactggct cccctgccc tgcccgcgcc atcggccgcc 15240
gtggagcggt cgcgtcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc 15300
gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360
caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac 15480
gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540
caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgtcgag 15600
ctgcgggccg acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc 15660
cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg 15720
atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780

atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgtcgctgct gcaccgcttc 15840
cgcgctcctgg accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900
gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960
tcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtacccgctc 16020
aagctggaaa ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080
gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg 16140
gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200
ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260
tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500
tggcattcgg cgctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560
acggcccca aa ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcgttgcc ggcggttta ttgctcgtga 16680
tgatcgtccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740
ttaatatttc gctattctgg agcttggtgt ttatttcggt ctaccgctg ccgggcgggg 16800
tcgcggcgac ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgctctgc 16860
taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg 16920
cgctgttggg gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980
cgggggcggt ttccatggcg ttcggaaccg tgctgaccgc caagtggcaa cctcccgtgc 17040

ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100

cttttagtggt tgatccgcc aatccgatgc ctacaggaac caatgttctc ggcctggcgt 17160

ggctcggcct gatcggagcg ggtttaacct acttccttg gttccggggg atctcgcgac 17220

tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280

cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcgggtgag 17340

caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400

ttcctcagcg gctttatcca gcgatttcct attatgtcgg catagttctc aagatcgaca 17460

gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520

agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580

ccgcgagatc atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc 17640

ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact 17700

gatgggctgc ctgtatcgag tgggtgatttt gtgccgagct gccggtcggg gagctgttgg 17760

ctggctgggtg gcaggatata ttgtgggtgta aacaaattga cgcttagaca acttaataac 17820

acattgcgga cgtttttaat gtactggggg ggtttttctt ttcaccagtg agacgggcaa 17880

cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg 17940

ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000

aatcaaaaag aatagcccga gatagggttg agtggtgttc cagtttgga caagagtcca 18060

ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120

ccactacgtg aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta 18180

aatcgggaacc ctaaaggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240

gcgagaaagg aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300

ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtaccggg 18449

<210> 41

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 41

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60
cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctgggtaag atccacgtct 240
tcgggaagcc agcgactggg gacctccagc gtccctttaa ggctgccaac agctttctca 300

gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gtccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttgatttg ctttccaggc tgagactcta gcttgagca tagagggtcc tttggcttcc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg	900
cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt	960
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa	1080
ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg	1140
gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct	1200
ggagtgacct agagggtcac gacttgagcc taaaatccgc cgctccacc atttgtagaa	1260
aatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga	1320
gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg	1380
ctctgaggtg cagtggatga ttattaatcc gggaccggcc gcccctccgc cccgaagtgg	1440
aaaggctggt gtgccccctg ttgaccaaga atctattgca tcatcggaga atatggagct	1500
tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg	1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctg ccgcccggtg tgcggcggg gttgacaagg tcgttgcgtc 1920

agtccaacat ttgttgccat attttcctgc tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc tcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct acccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctggaggggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcgctg gatatgtcct gggggtaaat agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcgcccgcg ctcccgattc cggaagtgct tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgccgtg cacagggtgt cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggg catggatgcg 2520

atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640

tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcaacgcgga tttcggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820

ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgcagg atcgccgcgg 2940
ctccggggcgt atatgctccg cattggtctt gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggg cgatgcgacg caatcgctccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgcccg cagaagcgcgg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcgctccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttgggtg cgatgtcagc tccggagttg agacaaatgg 3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagtg 3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaacctg gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggetcc acgcgactat 3600
atatttgtct ctaattgtac tttgacatgc tcctcttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720
ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780
gatggtatac aatagtaacc atgcatgggt gcctagttaa tgctccgtaa caccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020

agaaagtgaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080

ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140

ggagtttgaa attttgtcct ccaggagcaa aaaaaagttt ttttttatac atgtttgtac 4200

acaagaatag ttaccaattht gctttggctt tacgtgctgc aagtttatat cgttttcaat 4260

ttctttgtct ttacattttc tttgtccttt atctttcctc atttagtctt tgggagaatt 4320

aggaaaagg agcggaaagg taagaaatgc ttgcgtattht tactaattcg gcaaacatcc 4380

aatttggtcaa acagcagcct gtgcaacgct ctcgagatga cagtatcttht gattacactc 4440

taaatctcga tgacccgacc aaaaagagcg aacaaagaaa taatcttgtg cattogaata 4500

tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620

gctgcagtca atcagcgtca aggcccgccg cgttgaacta gcccgcgaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcgggtgctc gttagtctcg ctgagagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaacggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgag tgccatgggg 4980

tgaagtggct ggcactctcc tcttggtggt tgggtggcgcg ctccggcatgg agatgtatgc 5040

ccgctatgca caciaagcca tctggcatga gtcgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctcgactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgcccgcc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tggggggcggc 5220

ctgcttttga gcggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280

tggcctggtg cacaggcgct ttcccaccgg gcccatcgct ggccctgccct acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccaggtgcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctaggtgatc aaagactgtt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760

gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaaga ttgctctgtt. 5820

gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880

gattaagagt ttttttgttg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

aaagagacaa cacatacaat ataaaatcat tgtcttttaa tgtaattta gtagagtgt 6000

aagcctgcat tttttttgta cgcataaaca atgaattcac cccgcttctg gttttttaa 6060

aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttgttgt 6120

ggagtcacag gcttgtcttc agattgatag aggttgtata cactcaacag agcaatcttg 6180

gcacgttcgc ttcttttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt 6240

acagatccaa ttgcagcata tacttggcaa gcagctcgta caccgccttg acagtttgcc 6300

ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc 6360

ttcagagaca agcccaacag tctttgatca cctaggctac gggcgttgcc ttgctgtatt 6420

tgttctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480

gagccatact acagcataaa aggatacggt ttctttaaca gaaatttacc cttttgttat 6540

cagcacatac	aaaaaaaaag	aaattttaaga	tgagtaggac	ttccattctc	tcaaaaattt	6600
tattcaatcc	ataaatgaat	tattttttgga	caaaaaagaa	agattatgcc	tgatttttctc	6660
tattttttttt	tttttttaciaa	ctccaccaat	acttttctagc	ccagcttgge	gtaatcatgg	6720
tcatagctgt	ttcctgtgtg	aaattgttat	ccgctcacia	ttccacacia	catacgagcc	6780
ggaagcataa	agtgtaaagc	ctgggggtgcc	taatgagtga	gctaactcac	attaattgcg	6840
ttgcgctcac	tgcccgcctt	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	6900
ggccaacgcg	cggggagagg	cggtttgcgt	attgggccaa	agacaaaagg	gcgacattca	6960
accgattgag	ggaggggaagg	taaatattga	cggaatttat	tcattaaagg	tgaattatca	7020
ccgtcaccga	cttgagccat	ttgggaatta	gagccagcaa	aatcaccagt	agcaccatta	7080
ccattagcaa	ggccggaaac	gtcaccaatg	aaaccatcga	tagcagcacc	gtaatcagta	7140
gcgacagaat	caagtttgcc	tttagcgtca	gactgtagcg	cgttttcatc	ggcatttttcg	7200
gtcatagccc	ccttatttagc	gtttgccatc	ttttcataat	caaatcacc	ggaaccagag	7260
ccaccaccgg	aaccgcctcc	ctcagagccg	ccaccctcag	aaccgccacc	ctcagagcca	7320
ccaccctcag	agccgccacc	agaaccacca	ccagagccgc	cgccagcatt	gacaggaggc	7380
ccgatctagt	aacatagatg	acaccgcgcg	cgataattta	tcctagtttg	cgcgctatat	7440
tttgttttct	atcgcgattt	aatgtataa	ttgcgggact	ctaatcataa	aaaccatct	7500
cataaataac	gtcatgcatt	acatgttaat	tattacatgc	ttaacgtaat	tcaacagaaa	7560
ttatatgata	atcatcgcaa	gaccggcaac	aggattcaat	cttaagaaac	tttattgccca	7620
aatgtttgaa	cgatcgggga	tcacccgggt	ctgtggcggg	aactccacga	aaatatccga	7680
acgcagcaag	atatcgcggt	gcatctcggt	cttgccctggg	cagtcgccgc	cgacgcggtt	7740
gatgtggacg	ccgggcccga	tcattattgtc	gctcaggatc	gtggcgttgt	gcttgctggc	7800

cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc 7860

gaagaactcc agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920

gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980

gttgggctgc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040

agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100

aagcggtcag cccattcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg 8160

tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220

ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg 8280

tcgggcatgc gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340

togtccagat catcctgatc gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg 8400

cgatgtttcg cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460

attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520

tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580

acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc 8640

agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8700

gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcccc gtcatagccg 8760

aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820

cgaaacgac cagatccggg gcagattatt tggattgaga gtgaatatga gactctaatt 8880

ggataccgag gggaaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940

gctgatagtg accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000

agctcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060

gttccaaacg taaaacggct tgtcccgcgt catcgggcggg ggtcataacg tgactccctt 9120

aattctccgc tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 9180

acaggatata ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat 9240

ttaaaagggc gtgaaaaggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgcccgc ccgaaacgat 9360

ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420

gaatgccata gtgggcgggtg acgtcgttcg agtgaaccag atcgcgccagg aggcccggca 9480

gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgctc agaattacga 9540

tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600

cgcgccgatt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660

gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720

cgaggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780

gccggcgctt tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc 9840

catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt 9900

tctgatcggg aatgcccgca gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg 9960

catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020

cttcctctgc gagggcgggtt tttcggccgg ggacgccgtc aatgcgctga tgacaatcag 10080

ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140

cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggccgcga tagacgcctt 10200

cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260

ggcgaaaagg aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgac 10320

aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380

ccccctttcc acccggtcag acgcccgtag cagcccgcta cgggcttttt catgccctgc 10440

cctagcgtec aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500

gcttcctcgc tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 10560

cactcaaagg cggtaatagc gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620

tgagcaaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 10680

cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aaccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800

cctgttcoga ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaagcgtg 10860

gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggtatatcc 10920

atcctttttc gcacgatata caggattttg ccaaagggtt cgtgtagact ttccttggtg 10980

tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc 11040

cttcttcaact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100

ctggccgggt accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160

aaccaggaag ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220

tgaggaaaag gcggcgggcg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280

gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340

tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 11400

ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460

cgagcttggc aaggtcatga tgggcgtggt ccgcccgagg gcagagccat gactttttta 11520

gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580
agaagagcga cttcgcgagg ctggtgaagt acatcaccca cgagcaaggc aagaccgagc 11640
gcctttgcga cgctcaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc 11700
cctgcaaacg cgccagaaac gccgtcgaag ccgtgtgcga gacaccgcgg ccgccggcgt 11760
tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820
tgaggggccc actcacccgg cgcggcggtt acagatgagg ggcaggctcg atttcggccc 11880
gcgacgtgga gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc 11940
ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 12000
gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060
tgaggggcgc acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 12120
aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 12180
atatttataa accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgcgc 12240
aaggggggtg ccccccttc tcgaaccctc ccggcccgtc aacgcggggc tcccatcccc 12300
ccaggggctg cgcccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt 12360
ccttgccatt gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc 12420
cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480
tgagggcggc ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540
cttcatggcg gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600
cgtgctcgtg ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt 12660
ataccgaggt atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat 12720
ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ccttgaatat 12780

attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840

tttcaggggg caagggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca 12900

taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960

ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgcc 13020

gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080

tgccaggtgc tgccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140

gattacgtgc agctttccct tcaggcgga ttcatacagc ggccagccat ccgtcatcca 13200

tatcaccacg tcaaagggtg acagcaggct cataagacgc cccagcgtcg ccatagtgcg 13260

ttcaccgaat acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca 13320

gcgctggcgc gatthagccc cgacatagcc cactgttcg tccatttccg cgcagacgat 13380

gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggctgagtt ttttaagtga 13440

cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca 13500

ttcatggcca tatcaatgat tttctggtgc gtaccgggtt gagaagcggg gtaagtgaac 13560

tgcagttgcc atgttttacg gcagtgagag cagagatagc gctgatgtcc ggcggtgctt 13620

ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680

gccactggag cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc 13740

cataattgtg gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac 13800

aactttgaaa aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg 13860

gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920

ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaactg atcgaaaaat 13980

accgctgcgt aaaagatagc gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040

aaaatgaaaa cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg 14100

tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc 14160

tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcggtcc 14220

tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280

agtgcacatcag gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag 14340

acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg 14400

aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460

cgaaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520

ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580

agtggatatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640

atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700

atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760

caggagcgca ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820

aagtatttgg gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac 14880

gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940

gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc 15000

ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060

gaactgatcg acgcgggggt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120

atgcgtgcgc cccgcgaaac cttccagtcc gtcggctcga tgggtccagca agctacggcc 15180

aagatcgagc gcgacagcgt gcaactggct ccccctgccc tgcccgcgcc atcggccgcc 15240

gtggagcgtt cgcgtcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc 15300

gacacgcgag gaactatgac gaccaagaag cgaaaaaccg cggcgagga cctggcaaaa 15360
caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac 15480
gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540
caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcgag 15600
ctgcggggccg acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgcacc 15660
cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg 15720
atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780
atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgctgctgct gcaccgcttc 15840
cgcgtcctgg accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900
gtcgtgctgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960
tcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgctc 16020
aagctggaaa ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080
gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg 16140
gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200
ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260
tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500
tggcattcgg cgctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560

acggcccca aa ggacgctcac aaggcgcatc tgtccggcgt ttcgtggag cccgaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga 16680
tgatcgtccg acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac 16740
ttaatatctc gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg 16800
tcgcggcgac ggtaggcgct gtgcagccgc tgatggcgt gttcatctct gccgctctgc 16860
taggtagccc gatacgattg atggcgggtcc tgggggctat ttgcggaact gcgggcgtgg 16920
cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgctcgca gcgggcctgg 16980
cggggggcgt ttccatggcg ttcggaaccg tgctgaccgc caagtggcaa cctcccgtgc 17040
ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggccctggcgt 17160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220
tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280
cgggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag 17340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400
ttcctcagcg gctttatcca gcgatttcct attatgtcgg catagtctc aagatcgaca 17460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580
ccgcgagatc atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc 17640
ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact 17700
gatgggctgc ctgtatcgag tgggtatgtt gtgccgagct gccggtcggg gagctgttgg 17760
ctggctggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820

acattgcgga cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggt 17940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000
aaatcaaaag aatagcccga gatagggttg agtggtgttc cagtttggaa caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120
ccactacgtg aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta 18180
aatcggaacc ctaaaggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240
gcgagaaagg aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420
ggccagtgaa ttcgagctcg gtaccggg 18449

<210> 42

<211> 17593

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 42

ccgggctggt tgcctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgccct	480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaa cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc	660
tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgca ccccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg	840
ggtggcggcc tgcccttcac ttcgccgctc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140
agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaaa cttgcatgga 1260
ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgcgtgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtccgcatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggttca 2100
aatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttggtat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatTTTTTA aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggc attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatggtcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccc cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgcccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaagcg cgaggaatgc ctgtcgcgc tacaggcgac ggcatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgetgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcg 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttc 4380

gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgagcg agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatcgaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860
catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcggggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcatctt catcctcggc gcaacttaata tttcgctatt 5100
ctggagcttg ttgtttattht cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgt ctgctaggta gcccgatagc 5220
attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgccaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggcttttctca gccccagatc tggggtcgat cagccgggga tgcatacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggtttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacgggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggccacgc tggtttgccc cagcaggcga 6300

aaatcctggt tgatgggtgg tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

cggagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcata tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtgggtg caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440

tcgaaatata tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa ggcggttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcggtga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaadc cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggtcttgag gtgcagtgga tgattattaa tccgggaccg gccgccccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaacttatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640
tcagtccttg gtaggcagct ttgccccgtc tgtccgcccc gtgtgtcggc ggggttgaca 8700
aggctcgttg gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtag ttggtacat ccatactcca tccttcccat cccttattcc 8880
ttgaacctt tcagttcgag ctttcccact tcctcgcagc ttgactaaca gctaccccg 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtgctt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatgggtt 9120
ctacaaagat cgttatgttt atcggcactt tgcctcggcc gcgtccccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgtc gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgtc cgccgatct tagccagacg agcgggttcg gccattcgg 9360
accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgtgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggtc tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtag acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gtactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agtccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtcaa gcagcaaaga 10140

gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt ttttcgagtt 10800

tttttttttt ttctttgtga aggatttatt gttattggta tccatttttt attggaagac 10860

aagataagtt aatattgatt ttgcttaaag attaaaagga aatcagaaaa cgacaataaa 10920

aatgtaacg gacaaactat ggtgtcgatt ataagtctaa atccttaaaa aatgacaacg 10980

agttgctttc ctctgaaaac aattcttttg tctttgcaag aaaggtttct tttttgtttg 11040

cttgcatcac ttaaacaatca aatcaaatga aaggaataaa gcagatttga gggcgaataa 11100
ggattttctg gtcaacaaga tgtgagtgac acctaaggaa ctaaagcca ttcatttgtt 11160
ttaaaccgac atcaagatt gatgatcaac aggattgaga gagagaaaaa gaactcgtgt 11220
catttatttc tgttgactga aattttatat ttagaaaaaa tgtcaaactc atagcttttag 11280
ctatattaca taacatttga aataataata ataaaaaaag acacattaga gacacttttc 11340
aaactctaaa taactgtcta taaacacaaa gaaaacaaag acctctataa caacttatta 11400
gatttttctc gtacttttgt ctaaagatga tgtattcttg ttatcccaca cttctttcat 11460
ttgttcttga tgctactaaa tatacaaaaat ttcttttttg caagagatat tattccaaaa 11520
attttcaaaa agaaattttt ttcacaatag cagttgatcg tgtaacccaa agaggttctt 11580
tgttattttg cacttccgct ttgcggtgat gcatattcaa agtaatatat ggaataaaca 11640
acgtgtttta gcatgaaaga aaggaaacaa aggccgcttt gaacaaatgc ataattttc 11700
agacaaaaat gatctaaagc aagcagtaaa tcaacaaga aacattgctg attcgcgtta 11760
gaaaacgata aaagtcta atagccactaa gtatacttca atgaactttt tgtatgctta 11820
tggtccaatc agaccaataa tttgtgacca ttcttgaggt ggctttggtg atgcggaaac 11880
agaaaaaaat tttctacca atcgatttaa aaaacaattt ctgctttgaa ccaaaacttt 11940
ttttttctct ttaatcatta actttatcaa gtatgtacct acctcaaag tctcactca 12000
agcacaatta tgctaacatt gttccacctt ctcttttagaa atgttgtgga tttggaatgc 12060
cctgatcgtt ttcgttaaccg tgattggcat ggaagtgatt gctgcactgg cacacaaata 12120
catcatgcac ggctgggggtt ggggatggca tctttcacat catgaaccgc gtaaagggtgc 12180
gtttgaagtt aacgatcttt atgccgtggt ttttgctgca ttatcgatcc tgctgattta 12240
tctgggcagt acaggaatgt ggccgctcca gtggattggc gcaggatga cggcgtatgg 12300

attactctat tttatggtgc acgacgggct ggtgcatcaa cgttgccat tccgctatat 12360
tccacgcaag ggctacctca aacggttgta tatggcgcac cgtatgcatc acgccgtcag 12420
gggcaaagaa ggttggtgtt cttttggctt cctctatgcg ccgcccctgt caaaacttca 12480
ggcgacgctc cgggaaagac atggcgctag agcgggcgct gccagagatg cgcagggcgg 12540
ggaggatgag cccgcatccg ggaagtaagg gcctgaccag aggcggccag cagcagcggt 12600
aatTTTTcgg gcgtgggtcgt tgactgccgc tgatcccaaa gcttggcgta atcatgggtca 12660
tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc cacacaacat acgagccgga 12720
agcataaagt gtaaagcctg gggcgcctaa tgagttagct aactcacatt aattgcgttg 12780
cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 12840
caacgcgcgg ggagaggcgg tttgcgtatt gggccaaaga caaaaggcgg acattcaacc 12900
gattgagggg ggaaggttaa atattgacgg aaattattca ttaaagggtga attatcaccg 12960
tcaccgactt gagccatttg ggaattagag ccagcaaaat caccagtagc accattacca 13020
ttagcaaggc cggaaacgct accaatgaaa ccatcgatag cagcaccgta atcagtagcg 13080
acagaatcaa gtttgccttt agcgtcagac tgtagcgcggt tttcatcggc attttcggtc 13140
atagccccct tattagcggt tgccatcttt tcataatcaa aatcacccgga accagagcca 13200
ccaccggaac cgctccctc agagccgcca ccctcagaac cgccaccctc agagccacca 13260
ccctcagagc cgccaccaga accaccacca gagccgccc cagcattgac aggaggcccc 13320
atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgccg gctatatttt 13380
gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa cccatctcat 13440
aaataacgct atgcattaca tgttaattat tacatgctta acgtaattca acagaaatta 13500
tatgataatc atcgcaagac cggcaacagg attcaatctt aagaaacttt attgccaaat 13560

gtttgaacga tcggggatca tccgggtctg tggcgggaac tccacgaaaa tatccgaacg 13620
cagcaagata tcgcggtgca tctcgggtctt gcctgggcag tcgccgccga cgccggtgat 13680
gtggacgccg ggcccgatca tattgtcgct caggatcgtg gcgttggtgct tgtcggccgt 13740
tgctgtcgta atgatatcgg caccttcgac cgctgttcc gcagagatcc cgtgggcgaa 13800
gaactccagc atgagatccc cgcgctggag gatcatccag ccggcgctccc ggaaaacgat 13860
tccgaagccc aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcagggt 13920
gggcgtcgct tggtcgggtca tttcgaaccc cagagtccc ctcagaagaa ctcgtaaga 13980
aggcgataga aggcgatgcy ctgcgaatcg ggagcggcga taccgtaaag cacgaggaag 14040
cggtcagccc attcgccgcc aagctcttca gcaatatcac gggtagccaa cgctatgtcc 14100
tgatagcggc cggccacacc cagccggcca cagtcgatga atccagaaaa gcggccattt 14160
tccaccatga tattcggcaa gcaggcatcg ccatgggtca cgacgagatc atcgccgtcg 14220
ggcatgcgcg ccttgagcct ggcgaaacagt tcggctggcg cgagcccctg atgctcttcg 14280
tccagatcat cctgatcgac aagaccggct tccatccgag tacgtgctcg ctcgatgcga 14340
tgtttcgctt ggtggtcgaa tgggcaggta gccggatcaa gcgtatgcag ccgccgcatt 14400
gcatcagcca tgatggatac tttctcggca ggagcaaggc gagatgacag gagatcctgc 14460
cccggcactt cgcccaatag cagccagtcc cttcccgtt cagtgacaac gtcgagcaca 14520
gctgcgcaag gaacgcccgt cgtggccagc cacgatagcc gcgctgcctc gtccctgcagt 14580
tcattcaggc caccggacag gtcggtcttg aaaaaagaa ccgggcgccc ctgcgctgac 14640
agccggaaca cggcggcatc agagcagccg attgtctgtt gtgcccagtc atagccgaat 14700
agcctctcca cccaagcggc cggagaacct gcgtgcaatc catcttggtc aatcatgcga 14760
aacgatccag atccggtgca gattatttgg attgagagtg aatatgagac tctaattgga 14820

taccgagggg aatttatgga acgtcagtgg agcatttttg acaagaaata tttgctagct 14880
gatagtgacc ttaggcgact tttgaacgcg caataatggg ttctgacgta tgtgcttagc 14940
tcattaaact ccagaaaccc gcggctgagt ggctccttca acgttgcggt tctgtcagtt 15000
ccaaacgtaa aacggcttgt cccgcgtcat cggcgggggt cataacgtga ctcccttaat 15060
tctccgctca tgatcagatt gtcgtttccc gccttcagtt taaactatca gtgtttgaca 15120
ggatatattg gcgggtaaac ctaagagaaa agagcgttta ttagaataat cggatattta 15180
aaagggcgtg aaaaggttta tccgttcgtc catttgatg tgcatgccaa ccacagggtt 15240
ccccagatct ggcgccggcc agcgagacga gcaagattgg ccgccgcccg aaacgatccg 15300
acagcgcgcc cagcacaggt gcgcaggcaa attgcaccaa cgcatacagc gccagcagaa 15360
tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc gcgcaggagg cccggcagca 15420
ccggcataat caggccgatg ccgacagcgt cgagcgcgac agtgctcaga attacgatca 15480
ggggtatggt gggtttcacg tctggcctcc ggaccagcct ccgctgggtcc gattgaacgc 15540
gcggattcct tatcactgat aagttgggtg acatattatg tttatcagt ataaagtgtc 15600
aagcatgaca aagttgcagc cgaatacagt gatccgtgcc gccctggacc tggtgaacga 15660
ggtcggcgta gacggtctga cgacacgcaa actggcgga cggttggggg ttcagcagcc 15720
ggcgctttac tggcacttca ggaacaagcg ggcgtgctc gacgcactgg ccgaagccat 15780
gctggcggag aatcatagc attcgggtgcc gagagccgac gacgactggc gctcatttct 15840
gatcgggaat gcccgagct tcaggcaggc gctgctcgcc taccgcgatg gcgcgcgcat 15900
ccatgccggc acgcgaccgg gcgcaccgca gatggaaacg gccgacgcgc agcttcgctt 15960
cctctgcgag gcgggttttt cggccgggga cgccgtcaat gcgctgatga caatcagcta 16020
cttcactgtt ggggcogtgc ttgaggagca ggccggcgac agcgatgccg gcgagcgcg 16080

cggcaccgtt gaacaggctc cgctctcgcc gctgttgccg gccgcgatag acgccttcga 16140
cgaagccggg ccggacgcag cgttcgagca gggactcgcg gtgattgtcg atggattggc 16200
gaaaaggagg ctcgttgtca ggaacgttga aggaccgaga aagggtgacg attgatcagg 16260
accgctgccg gagcgcaacc cactcactac agcagagcca tgtagacaac atccccctccc 16320
cctttccacc gcgtcagacg cccgtagcag cccgctacgg gctttttcat gccctgccct 16380
agcgtccaag cctcacggcc gcgctcgccc tctctggcgg ccttctggcg ctcttcgct 16440
tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggg atcagctcac 16500
tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga 16560
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 16620
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 16680
ccgacaggac tataaagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct 16740
gttccgacct tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg 16800
cttttcgct gcataaccct gcttcggggg cattatagcg attttttcgg tatatccatc 16860
ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc cttggtgtat 16920
ccaacggcgt cagccgggca ggataggtga agtaggcccc cccgcgagcg ggtgttcctt 16980
cttcaactgtc ccttatctgc acctggcggt gctcaacggg aatcctgctc tgcgaggctg 17040
gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa ccaagccaac 17100
caggaagggc agcccaccta tcaaggtgta ctgccttcca gacgaacgaa gagcgattga 17160
ggaaaaggcg gcggcgcccg gcatgagcct gtcggcctac ctgctggccg tcggccaggg 17220
ctacaaaatc acgggcgtcg tggactatga gcacgtccgc gagctggccc gcatcaatgg 17280
cgacctgggc cgctggggcg gcctgctgaa actctggctc accgacgacc cgcgcacggc 17340

gcgggttcggt gatgccacga tcctcgccct gctggcgaag atcgaagaga agcaggacga 17400
gcttggcaag gtcatgatgg gcgtgggtccg cccgagggca gagccatgac ttttttagcc 17460
gctaaaacgg ccgggggggtg cgcgtgattg ccaagcacgt ccccatgcgc tccatcaaga 17520
agagcgactt cgcggagctg gtgaagtaca tcaccgacga gcaaggcaag accgagcgcc 17580
tttgcgacgc tca 17593

<210> 43
<211> 16954
<212> DNA
<213> Artificial

<220>
<223> Plasmid

<220>
<221> misc_feature
<222> (10264)..(10264)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10472)..(10472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (10563)..(10563)
<223> n is a, c, g, or t

<400> 43
ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtgga tacctcgcg 120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180

ccggcgcggc gttgacagat gaggggacagg ctcgatttcg gccggcgacg tggagctggc 240
cagcctcgca aatcgggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga 300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaaggg ttcgcccgt 480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540
tttttaacca gggctgcgcc ctgtgcgct gaccgcgcac gccgaaggg ggtgcccccc 600
cttctcgaac cctcccggcc cgctaacggg ggccctccat cccccaggg gctgcgcccc 660
tcggccgcga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720
atcggggcag taacgggatg ggcgatcagc ccgagcgga cgcgcggaag cattgacgtg 780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840
ggtggcgggc tgccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccc 900
gcaattttta ccttgggcat tcttggcata gtggtcgagg gtgccgtgct cgtgttcggg 960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140
agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260
ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttcggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgctccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgogg gctgttgccc ggcattcaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220
aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatatTT tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggcttacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg ttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgtctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaagge cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgagge agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatgtcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gtcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgtatt 5100

ctggagcttg ttgtttatct cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtcctggggg ctatttgccg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgcaa ctggcgggcg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctcgc agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgcggtccc gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccc agctgccggg cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300
aaatcctgtt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
caccctaaatc aagttttttg gggctcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600
agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcgggtacc cgggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcatc tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgtga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200
cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttggga gcatagaggg 7320
tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440
tcgaaataca tccgatgtc gaaggcttgg ggcacctgcg ttgggtgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920
tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980
cctagctgat tctggagtga cccagaggggt catgacttga gcctaaaate cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tggtcatcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaagggt ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640
tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtcaca catttgttgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcccct aagtaagtac ttgctacat ccatactcca tcttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgtg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgcgcaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agtccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tatttttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacggcagc tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800

tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860

atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920

aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt ttttttttat 10980

acatgtttgt acacaagaat agttaccaat ttgctttggc cttacgtgct gcaagtttat 11040

atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100

tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160

cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220

ttgattacac tctaaatctc gatgacccga caaaaagag cgaacaaaga aataatcttg 11280

tgcatcgaat tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400

tatgttgtgg atttggaaatg ccctgatcgt tttcgttacc gtgattggca tggaagtgat 11460
tgctgcactg gcacacaaat acatcatgca cggctgggggt tggggatggc atctttcaca 11520
tcatgaaccg cgtaaagggtg cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580
attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640
cgcaggtatg acggcgtatg gattactcta ttttatgggtg cacgacgggc tgggtgcatca 11700
acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760
ccgtatgcat cacgccgtca ggggcaaaga aggttgtgtt tcttttggct tcctctatgc 11820
gccgcccctg tcaaaacttc aggcgacgt cgggaaaga catggcgcta gagcgggcgc 11880
tgccagagat gcgcaggcg gggaggatga gcccgcatcc gggaagtaag ggcctgacca 11940
gaggcggcca gcagcagcgt taatttttcg ggcgtggtcg ttgactgccg ctgatcccaa 12000
agcttggcgt aatcatggtc atagctgtt cctgtgtgaa attgttatcc gctcacaatt 12060
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 12120
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 12180
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcaaaag 12240
acaaaagggc gacattcaac cgattgaggg agggaaggta aatattgacg gaaattattc 12300
attaaagggtg aattatcacc gtcaccgact tgagccattt gggaattaga gccagcaaaa 12360
tcaccagtag caccattacc attagcaagg ccggaaacgt caccaatgaa accatcgata 12420
gcagcaccgt aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480
ttttcatcgg cattttcgggt catagcccc ttattagcgt ttgccatctt ttcataatca 12540
aatcaccgg aaccagagcc accaccggaa ccgcctccct cagagccgcc accctcagaa 12600
ccgccacct cagagccacc accctcagag ccgccaccag aaccaccacc agagccgccg 12660

ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataatttatc 12720
ctagtttgcg cgctatattt tgttttctat cgcgtattaa atgtataatt gcgggactct 12780
aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 12840
aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900
taagaaactt tattgcaaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 12960
ctccacgaaa atatccgaac gcagcaagat atcgcggtgc atctcggtct tgcctgggca 13020
gtcgccgcgcg acgccgttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080
ggcgttgtgc ttgtcgcccg ttgctgtcgt aatgatatcg gcaccttcga ccgcctgttc 13140
cgcagagatc ccgtgggcga agaactccag catgagatcc ccgcgctgga ggatcatcca 13200
gccggcgctcc cggaaaacga ttccgaagcc caacctttca tagaaggcgg cgttggaatc 13260
gaaatctcgt gatggcaggt tgggcgtcgc ttggtcggtc atttcgaacc ccagagtccc 13320
gctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg 13380
ataccgtaaa gcacgaggaa gcggtcagcc cattcgccgc caagctcttc agcaatatca 13440
cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc acagtcgatg 13500
aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560
acgacgagat catcgccgtc gggcatgcgc gccttgagcc tggcgaacag ttcggctggc 13620
gcgagcccct gatgctcttc gtccagatca tctgatcga caagaccggc ttccatccga 13680
gtacgtgctc gctcgatgcg atgtttcgct tggtggtcga atgggcaggt agccggatca 13740
agcgtatgca gccgccgcat tgcattcagcc atgatggata ctttctcggc aggagcaagg 13800
tgagatgaca ggagatcctg ccccggcact tcgcccaata gcagccagtc ccttcccgtc 13860
tcagtgacaa cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc 13920

cgcgctgcct cgtcctgcag ttcattcagg gcaccggaca ggtcgggtctt gacaaaaaga 13980
accgggcgcc cctgcgctga cagccggaac acggcggcat cagagcagcc gattgtctgt 14040
tgtgcccagt catagccgaa tagcctctcc acccaagcgg cgggagaacc tgcgtgcaat 14100
ccatcttggt caatcatgcg aaacgatcca gatccgggtgc agattatttg gattgagagt 14160
gaatatgaga ctctaattgg ataccgaggg gaatttatgg aacgtcagtg gagcattttt 14220
gacaagaaat atttgctagc tgatagtgac cttaggcgac ttttgaacgc gcaataatgg 14280
tttctgacgt atgtgcttag ctcatataac tccagaaacc cgcgggtgag tggctccttc 14340
aacgttgagg ttctgtcagt tccaaacgta aaacgggttg tcccgcgta tcggcggggg 14400
tcataacgtg actcccttaa ttctccgctc atgacagat tgcgttttc cgccttcagt 14460
ttaaactatc agtgtttgac aggatatatt ggcgggtaaa cctaagagaa aagagcgttt 14520
attagaataa tcggatattt aaaagggcgt gaaaagggtt atccgttcgt ccatttgtat 14580
gtgcatgcca accacagggg tccccagatc tggcgccggc cagcgagacg agcaagattg 14640
gccgccgcc gaaacgatcc gacagcgcg ccagcacagg tgcgcaggca aattgcacca 14700
acgcatacag cgccagcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760
cgcgaggag gcccggcagc accggcataa tcaggccgat gccgacagcg tcgagcgcg 14820
cagtgtcag aattacgac aggggtatgt tgggtttcac gtctggcctc cggaccagcc 14880
tccgctggc cgattgaacg cgcggattct ttatcactga taagttggtg gacatattat 14940
gtttatcagt gataaagtgt caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000
cgccctggac ctgttgaacg aggtcggcgt agacggtctg acgacacgca aactggcgga 15060
acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120
cgacgcactg gccgaagcca tgctggcgga gaatcatag cattcgggtgc cgagagccga 15180

cgacgactgg cgctcatttc tgatcgggaa tgcccgcagc ttcaggcagg cgctgctcgc 15240

ctaccgcgat ggcgcgcgca tccatgccgg cacgcgaccg ggcgcaccgc agatggaaac 15300

ggccgacgcg cagcttcgct tcctctgcga ggccgggtttt tcggccgggg acgccgtcaa 15360

tgcgctgatg acaatcagct acttcactgt tggggccgtg cttgaggagc aggccggcga 15420

cagcgatgcc ggcgagcgcg gcggcaccgt tgaacaggct ccgctctcgc cgctgttgcg 15480

ggccgcgata gacgccttcg acgaagccgg tccggacgca gcgttcgagc agggactcgc 15540

ggtgattgtc gatggattgg cgaaaaggag gctcgttgtc aggaacgttg aaggaccgag 15600

aaagggtgac gattgatcag gaccgctgcc ggagcgcaac ccactacta cagcagagcc 15660

atgtagacaa catcccctcc ccctttccac cgcgctcagac gcccgtagca gcccgctacg 15720

ggctttttca tgccctgccc tagcgccaa gcctcacggc cgcgctcggc ctctctggcg 15780

gccttctggc gctcttcgcg ttccctcgctc actgactcgc tgcgctcggg cgcttcggctg 15840

cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat 15900

aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 15960

gcgttgctgg cgttttttcca taggctccgc cccctgacg agcatcacia aaatcgacgc 16020

tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 16080

agctccctcg tgcgctctcc tgttcgcgacc ctgccgctta ccggatacct gtccgccttt 16140

ctcccttcgg gaagcgtggc gctttttcgc tgcataaccc tgcttcgggg tcattatagc 16200

gattttttcg gtatatccat cctttttcgc acgatataca ggattttgcc aaaggggttcg 16260

tgtagacttt ccttggtgta tccaaaggcg tcagccgggc aggatagggtg aagtaggccc 16320

accgcgagc ggggtgttcc tcttcactgt cccttattcg cacctggcgg tgcacaacgg 16380

gaatcctgct ctgcgaggct ggccggctac cgccggcgta acagatgagg gcaagcggat 16440

ggctgatgaa accaagccaa ccaggaaggg cagcccacct atcaagggtgt actgccttcc 16500
agacgaacga agagcgattg aggaaaaggc ggcggcggcc ggcattgagcc tgcgggccta 16560
cctgctggcc gtcggccagg gctacaaaat cacgggcgtc gtggactatg agcacgtccg 16620
cgagctggcc cgcattcaatg gcgacctggg ccgcctgggc ggcctgctga aactctggct 16680
caccgacgac ccgcgcacgg cgcgggttcgg tgatgccacg atcctcgccc tgctggcgaa 16740
gatcgaagag aagcaggacg agcttggtcaa ggtcatgatg ggcgtggtcc gcccgagggc 16800
agagccatga ctttttttagc cgctaaaacg gccgggggggt gcgcgtgatt gccaagcacg 16860
tccccatgcg ctccatcaag aagagcgact tcgcggagct ggtgaagtac atcaccgacg 16920
agcaaggcaa gaccgagcgc ctttgcgacg ctca 16954

<210> 44

<211> 16954

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 44

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtgga tacctcgccg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttcccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatatth ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaa cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc	660
tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgca cggccggaag cattgacgtg	780
ccgcagggtc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg	840
ggtggcggcc tgcccttcac ttccggcgtc ggggcattca cggacttcat ggcggggccg	900
gcaatthtta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atattthaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtgcgcatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgctccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattccaa gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100

aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagtgc gtcttggtat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360

gagggatgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggt gtggcagcag gtggtggagt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccgggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtacct gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatgtcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcg tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860

catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggtatgctg ctgcgggctg tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttatct cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtccctggggg ctatttgctg aactgcgggc gtggcgctgt tgggtgtgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgga caatggcgagg gaggacttct gctcgttcca gtagcttttag tgtttgatcc 5460
gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgaactcgaac ctacagttgt 5580
ttccttactg ggcttttctc gcccagatc tggggtcgat cagccgggga tgcatcaggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgct aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgct tacaacggct ctcccgtga cgccgtccc gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctggt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gtcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggtgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcggttgc caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaadc cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggtctctgag gtgcagtgga tgattattaa tccgggacct gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtcacctg gtaggcagct ttgccccgtc tgtccgcccc gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac tttgtacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggogaagaat ctcggtgcttt 9060

cagcttcgat gtaggggggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacgggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtaacgttg caagacctgc ctgaaaccga actgcccgtc gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgtg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgtgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgcgcaggg 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggtc tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttggggcgag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgcctctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttcttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg ttaccacaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctctgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800
taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860
tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920
ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca aagatactgt 10980
catctcgaga gcgttgacaca ggctgctgtt tgccaaattg gatgtttgcc gaattagtaa 11040
aatacgcaag catttcttac ctttccgctc ctttttcta attctcccaa agactaaatg 11100
aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160

cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaacatg tataaaaaaa 11220

aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgata 11280

tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta ttataaaaa 11340

gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct attttaatct 11400

catgttgtgg atttggaatg ccctgatcgt tttcgttacc gtgattggca tggaagtgat 11460

tgctgcactg gcacacaaat acatcatgca cggctgggggt tggggatggc atctttcaca 11520

tcatgaaccg cgtaaagggt cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580

attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640

cgcaggatat acggcgtatg gattactcta ttttatgggt cacgacgggc tggtgcatca 11700

acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760

cgtatgcat cacgccgtca ggggcaaaga aggttgtgtt tcttttggct tcctctatgc 11820

gccgcccctg taaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcgggcgc 11880

tgccagagat gcgcaggcg gggaggatga gcccgcatcc gggaagtaag ggcctgacca 11940

gaggcggcca gcagcagcgt taatttttcg ggcgtggtcg ttgactgccg ctgatcccaa 12000

agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gtcacaatt 12060

ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 12120

taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 12180

cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcaaaag 12240

acaaaagggc gacattcaac cgattgaggg agggaaggta aatattgacg gaaattattc 12300

attaaaggtg aattatcacc gtcaccgact tgagccattt gggaattaga gccagcaaaa 12360

tcaccagtag caccattacc attagcaagg ccggaaacgt caccaatgaa accatcgata 12420

gcagcaccgt aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12480

ttttcatcgg cattttcggc catagccccc ttattagcgt ttgccatctt ttcataatca 12540

aaatcaccgg aaccagagcc accaccggaa ccgcctccct cagagccgcc accctcagaa 12600

ccgccaccct cagagccacc accctcagag ccgccaccag aaccaccacc agagccgccg 12660

ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataatttata 12720

ctagtttgcg cgctatattt tgttttctat cgcgtattaa atgtataatt gcgggactct 12780

aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 12840

aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 12900

taagaaactt tattgccaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 12960

ctccacgaaa atatccgaac gcagcaagat atcgcggtgc atctcggtct tgcctgggca 13020

gtcgccgccg acgccgttga tgtggacgcc gggcccgatc atattgtcgc tcaggatcgt 13080

ggcgttgtgc ttgtcgcccg ttgctgtcgt aatgatatac gcaccttcga ccgcctgttc 13140

cgcagagatc ccgtgggcga agaactccag catgagatcc ccgcgctgga ggatcatcca 13200

gccggcgctc cggaaaacga ttccgaagcc caacctttca tagaaggcgg cggtggaatc 13260

gaaatctcgt gatggcaggt tgggcgtcgc ttggtcggtc atttcgaacc ccagagtccc 13320

gctcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg 13380

ataccgtaaa gcacgaggaa gcggtcagcc cattcgccgc caagctcttc agcaatatca 13440

cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc acagtcatg 13500

aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc 13560

acgacgagat catcgccgtc gggcatgcgc gccttgagcc tggcgaacag ttcggtggc 13620

gcgagcccct gatgctcttc gtccagatca tcctgatcga caagaccggc ttccatccga 13680

gtacgtgctc gctcgatgcg atgtttcgct tgggtggtcga atgggcaggt agccggatca 13740
agcgtatgca gccgccgcat tgcacagcc atgatggata ctttctcggc aggagcaagg 13800
tgagatgaca ggagatcctg ccccggcact tcgcccataa gcagccagtc ctttcccgtc 13860
tcagtgacaa cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc 13920
cgcgctgcct cgtcctgcag ttcattcagg gcaccggaca ggtcggctctt gacaaaaaga 13980
accgggcgcc cctgcgctga cagccggaac acggcggcat cagagcagcc gattgtctgt 14040
tgtgcccagt catagccgaa tagcctctcc acccaagcgg ccggagaacc tgcgtgcaat 14100
ccatcttggt caatcatgcg aaacgatcca gatccggtgc agattatttg gattgagagt 14160
gaatatgaga ctctaattgg ataccgaggg gaatttatgg aacgtcagtg gagcattttt 14220
gacaagaaat atttgctagc tgatagtgac cttaggcgac ttttgaacgc gcaataatgg 14280
tttctgacgt atgtgcttag ctcatataac tccagaaacc cgcggctgag tggctccttc 14340
aacgttgagg ttctgtcagt tccaaacgta aaacggcttg tcccgcgtca tcggcggggg 14400
tcataacgtg actcccttaa ttctccgctc atgatcagat tgcggtttcc cgccttcagt 14460
ttaaactatc agtgtttgac aggatatatt ggcgggtaaa cctaagagaa aagagcgttt 14520
attagaataa tcggatattt aaaagggcgt gaaaagggtt atccgttcgt ccatttgtat 14580
gtgcatgcca accacagggg tcccagatc tggcgccggc cagcgagacg agcaagattg 14640
gccgccgcc gaaacgatcc gacagcgcg ccagcacagg tgcgcaggca aattgcacca 14700
acgcatacag cgccagcaga atgccatagt gggcggtgac gtcgttcgag tgaaccagat 14760
cgcgcaggag gcccggcagc accggcataa tcaggccgat gccgacagcg tcgagcgcg 14820
cagtgtcag aattacgatc aggggtatgt tgggtttcac gtctggcctc cggaccagcc 14880
tccgctggtc cgattgaacg cgcggattct ttatcactga taagttgggtg gacatattat 14940

gtttatcagt gataaagtgt caagcatgac aaagttgcag ccgaatacag tgatccgtgc 15000
cgccctggac ctgttgaacg aggtcggcgt agacggtctg acgacacgca aactggcgga 15060
acggttgggg gttcagcagc cggcgcttta ctggcacttc aggaacaagc gggcgctgct 15120
cgacgcactg gccgaagcca tgctggcgga gaatcatacg cattcggtgc cgagagccga 15180
cgacgactgg cgctcatttc tgatcgggaa tgccgcagc ttcaggcagg cgctgctgc 15240
ctaccgcgat ggcgcgcgca tccatgccgg cagcgcaccg ggcgcaccgc agatggaaac 15300
ggccgacgcg cagcttcgct tcctctgcga ggcggtttt tcggccgggg acgccgtcaa 15360
tgcgctgatg acaatcagct acttactgt tggggccgtg cttgaggagc aggccggcga 15420
cagcgatgcc ggcgagcgcg gcggcaccgt tgaacaggct ccgctctcgc cgctgttgcg 15480
ggccgcgata gacgccttcg acgaagccgg tccggacgca gcgctcgagc agggactcgc 15540
ggtgattgtc gatggattgg cgaaaaggag gctcgttgtc aggaacgttg aaggaccgag 15600
aaagggtgac gattgatcag gaccgctgcc ggagcgcaac ccactcacta cagcagagcc 15660
atgtagacaa catccctcc ccctttccac cgcgtcagac gcccgtagca gcccgctacg 15720
ggctttttca tgccctgcc tagcgtccaa gcctcacggc cgcgctcggc ctctctggcg 15780
gccttctggc gctcttcgc ttctcgtc actgactcgc tgcgctcggt cgctcggctg 15840
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat 15900
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 15960
gcgttgctgg cgtttttcca taggctcgc cccctgacg agcatcacia aaatcgacgc 16020
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt tccccctgga 16080
agctccctcg tgcgctctcc tgttcgacc ctgccgctta ccggatacct gtccgccttt 16140

ctcccttcgg gaagcgtggc gcttttccgc tgcataaccc tgcttcgggg tcattatagc 16200
gatttttttcg gtatatccat cctttttcgc acgatataca ggattttgcc aaaggggttcg 16260
tgtagacttt ccttggtgta tccaacggcg tcagccgggc aggataggtg aagtaggccc 16320
acccgcgagc ggggtgttct tcttctactgt cccttattcg cacctggcgg tgctcaacgg 16380
gaatcctgct ctgcgaggct ggccggctac cgccggcgta acagatgagg gcaagcggat 16440
ggctgatgaa accaagccaa ccaggaaggc cagcccacct atcaagggtg actgccttcc 16500
agacgaacga agagcgattg aggaaaaggc ggcggcggcc ggcatgagcc tgtcggccta 16560
cctgctggcc gtcggccagg gctacaaaat cacgggcgtc gtggactatg agcacgtccg 16620
cgagctggcc cgcacatg ggcacctggg ccgcctgggc ggctgctga aactctggct 16680
caccgacgac ccgcgcacgg cgcgggttcgg tgatgccacg atcctcgccc tgetggcgaa 16740
gatcgaagag aagcaggacg agcttggtcaa ggtcatgatg ggcgtgggtcc gcccgagggc 16800
agagccatga ctttttttagc cgctaaaacg gccgggggggt gcgcgtgatt gccaagcacg 16860
tccccatgcg ctccatcaag aagagcgact tcgcggagct ggtgaagtac atcaccgacg 16920
agcaaggcaa gaccgagcgc ctttgcgacg ctca 16954

<210> 45

<211> 19491

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (18970)..(18970)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19178)..(19178)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19269)..(19269)

<223> n is a, c, g, or t

<400> 45

agcttggtac cgagctcgga tccactagta acggccgcca gtgtgctgga attcgccctt	60
gacggccagt gaattcgagc tcggtacccg gggatctttc gacactgaaa tacgtcgagc	120
ctgctccgct tggaagcggc gaggagcctc gtcctgtcac aactaccaac atggagtacg	180
ataagggcca gttccgccag ctcattaaga gccagttcat gggcggtggc atgatggccg	240
tcatgcatct gtacttcaag tacaccaacg ctcttctgat ccagtcgatc atccgctgaa	300
ggcgctttcg aatctggtta agatccacgt cttcgggaag ccagcgactg gtgacctcca	360
gcgtcccttt aaggctgcca acagctttct cagccagggc cagcccaaga ccgacaaggc	420
ctccctccag aacgccgaga agaactggag gggtggtgtc aaggaggagt aagctcctta	480
ttgaagtcgg aggacggagc ggtgtcaaga ggatattctt cgactctgta ttatagataa	540
gatgatgagg aattggagggt agcatagctt catttggatt tgctttccag gctgagactc	600
tagcttggag catagagggt cctttggctt tcaatattct caagtatctc gagtttgaac	660
ttattccctg tgaacctttt attcaccaat gagcattgga atgaacatga atctgaggac	720
tgcaatcgcc atgagggttt cgaaatacat ccggatgtcg aaggcttggg gcacctgcgt	780
tggttgaatt tagaacgtgg cactattgat catccgatag ctctgcaaag ggcgttgcac	840
aatgcaagtc aaacgttgct agcagttcca ggtggaatgt tatgatgagc attgtattaa	900

atcaggagat atagcatgat ctctagttag ctcaccacaa aagtcagacg gcgtaaccaa 960

aagtcacaca acacaagctg taaggatttc ggcacggcta cggaagacgg agaagccacc 1020

ttcagtggac tcgagtacca ttttaattcta tttgtgtttg atcgagacct aatacagccc 1080

ctacaacgac catcaaagtc gtatagctac cagtgaggaa gtggactcaa atcgacttca 1140

gcaacatctc ctggataaac ttttaagccta aactatacag aataagatag gtggagagct 1200

tataccgagc tcccaaattct gtccagatca tggttgaccg gtgcctggat cttcctatag 1260

aatcatcctt attcgttgac ctagctgatt ctggagtgac ccagaggggc atgacttgag 1320

cctaaaatcc gccgcctcca ccattttagt aaaaatgtga cgaactcgtg agctctgtac 1380

agtgaccggt gactctttct ggcattgcga gagacggacg gacgcagaga gaagggctga 1440

gtaataagcc actggccaga cagctctggc ggctctgagg tgcagtggat gattattaat 1500

ccgggaccgg ccgccccctc gccccgaagt ggaaaggctg gtgtgcccc cgttgaccaa 1560

gaatctattg catcatcgga gaatatggag cttcatcgaa tcaccggcag taagcgaagg 1620

agaatgtgaa gccaggggtg tatagccgtc ggcgaaatag catgccatta acctaggtac 1680

agaagtccaa ttgcttccga tctggtaaaa gattcacgag atagtacctt ctccgaagta 1740

ggtagagcga gtaccggcg cgtaagctcc ctaattggcc catccggcat ctgtagggcg 1800

tccaaatatc gtgcctctcc tgctttgccc ggtgtatgaa accggaaagg ccgctcagga 1860

gctggccagc ggcgcagacc gggaacacaa gctggcagtc gacccatccg gtgctctgca 1920

ctcgacctgc tgaggtcctt cagtccctgg taggcagctt tgccccgtct gtccgcccgg 1980

tgtgtcggcg gggttgacaa ggtcgttgcg tcagtccaac atttgttgcc atattttcct 2040

gtctccccca ccagctgctc ttttcttttc ttttctttt cccatcttca gtatattcat 2100

cttcccatcc aagaaccttt atttccccta agtaagtact ttgctacatc catactccat 2160

ccttcccatc ccttattcct ttgaaccttt cagttcgagc tttcccactt catcgagct 2220

tgactaacag ctaccccgct tgagcagaca tcaccatgct gtcgaagctg cagtcaatca 2280

gcgtcaaggc ccgccgcgtt gaactagccc gcgacatcac gcggcccaa gtctgcctgc 2340

atgctcagcg gtgctcgta gttcggctgc gagtggcagc accacagaca gaggaggcgc 2400

tgggaaccgt gcaggctgcc ggcgcgggcg atgagcacag cgccgatgta gactccagc 2460

agcttgaccg ggctatcgca gagcgctcg cccggcgcaa acgggagcag ctgtcatacc 2520

aggctgccgc cattgcagca tcaattggcg tgtcaggcat tgccatcttc gccacctacc 2580

tgagatttgc catgcacatg accgtgggcg gcgcagtgcc atggggtgaa gtggctggca 2640

ctctcctctt ggtggttggg ggcgcgctcg gcatggagat gtatgccgc tatgcacaca 2700

aagccatctg gcatgagtcg cctctgggct ggctgctgca caagagccac cacacacctc 2760

gactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc 2820

tcctgtgtac ctttggcttc tggctgcca acgtcctggg gcgggcctgc tttggagcgg 2880

ggctgggcat cagctatac ggcatggcat atatgtttgt acacgatggc ctggtgcaca 2940

ggcgctttcc caccgggccc atcgctggcc tgccctacat gaagcgctg acagtggccc 3000

accagctaca ccacagcggc aagtacggtg gcgcgccctg gggatatgtc ttgggtccac 3060

aggagctgca gcacattcca ggtgcggcgg aggaggtgga gcgactggtc ctggaactgg 3120

actggtccaa gcggtagggt gcggaaccag gcacgctggt ttcacacctc atgcctgtga 3180

taagggtgtg ctagagcgat gcgtgtgaga cgggtatgtc acggtcgact ggtctgatgg 3240

ccaatggcat cggccatgtc tggatcacac gggctgggtg cctgggtgaa ggtgatgcac 3300

atcatcatgt gcggttggag gggctggcac agtgtgggct gaactggagc agttgtccag 3360

gctggcggttg aatcagtgag ggtttgtgat tggcggttgt gaagcaatga ctccgccccat 3420

attctatttg tgggagctga gatgatggca tgcttgggat gtgcatggat catggtagtg 3480

cagcaaaacta tattcaccta gggctgttgg taggatcagg tgaggccttg cacattgcat 3540

gatgtactcg tcatgggtgtg ttggtgagag gatggatgtg gatggatgtg tattctcaga 3600

cgtagacctt gactggagggc ttgatcgaga gagtgggccc tattctttga gaggggagggc 3660

tcgtgccaga aatgggtgagt ggatgactgt gacgctgtac attgcaggca ggtgagatgc 3720

actgtctcga ttgtaaaata cattcagatg caagcttggc gtaatcatgg tcatagctgt 3780

ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacgagcc ggaagcataa 3840

agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac 3900

tgcccgcctt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 3960

cggggagagg cggtttgcgt attggggcaa agacaaaagg ggcacattca accgattgag 4020

ggaggggaagg taaatattga cggaaattat tcattaaagg tgaattatca ccgtcaccga 4080

cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4140

ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat 4200

caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg gtcatagccc 4260

ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg 4320

aaccgcctcc ctgagagccg ccaccctcag aaccgccacc ctgagagcca ccaccctcag 4380

agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggagggc ccgatctagt 4440

aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct 4500

atcgcgattt aaatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac 4560

gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4620

atcatcgcaa gaccggcaac aggattcaat ctttaagaaac tttattgcca aatgtttgaa 4680

cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag 4740

atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg 4800

cggggcccgga tcatattgtc gctcaggatc gtggcggttg gcttgtcggc cgttgctgtc 4860

gtaatgatat cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc gaagaactcc 4920

agcatgagat ccccgcgctg gaggatcacc cagccggcgt cccggaaaac gattccgaag 4980

cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag gttgggcgtc 5040

gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca agaaggcgat 5100

agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag 5160

cccatcgcg gccaaagctct tcagcaatat cacgggtagc caacgctatg tcctgatagc 5220

ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca ttttccacca 5280

tgatatcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc 5340

gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat 5400

catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg 5460

cttgggtggc gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag 5520

ccatgatgga tacttttctc gcaggagcaa ggtgagatga caggagatcc tgccccggca 5580

cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc 5640

aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctgctcctgc agttcattca 5700

gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga 5760

acacggcggc atcagagcag ccgattgtct gttgtgccca gtcatagccg aatagcctct 5820

ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgac 5880

cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag 5940

gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg 6000

accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa 6060

actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg 6120

taaaacggct tgtcccgcgt catcggcggg ggtcataacg tgactccctt aattctccgc 6180

tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg acaggatata 6240

ttggcgggta aacctagag aaaagagcgt ttattagaat aatcggatat ttaaagggc 6300

gtgaaaaggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga 6360

tctggcgccg gccagcgaga cgagcaagat tggccgccgc ccgaaacgat ccgacagcgc 6420

gcccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca gaatgccata 6480

gtgggcgggtg acgtcgttcg agtgaaccag atcgcgcagg agggccggca gcaccggcat 6540

aatcaggccg atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat 6600

gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggtt 6660

ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg 6720

acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggctggc 6780

gtagacggtc tgacgacacg caaactggcg gaacgggttg gggttcagca gccggcgctt 6840

tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg 6900

gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt tctgatcggg 6960

aatgcccgcg gcttcaggca ggcgctgtc gcctaccgcg atggcgcgcg catccatgcc 7020

ggcacgcgac cgggcgaccc gcagatggaa acggccgacg cgcagcttcg ctctctctgc 7080

gaggcggtt tttcgccggg ggacgccgtc aatgcgctga tgacaatcag ctacttcact 7140

gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc 7200

gttgaacagg ctccgctctc gccgctgttg cgggccgcga tagacgcctt cgacgaagcc 7260

ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg 7320

aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc aggaccgctg 7380

ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct cccctttcc 7440

accgcgtcag acgcccgtag cagcccgtc cgggcttttt catgccctgc cctagcgtcc 7500

aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttcctcgc 7560

tactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 7620

cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 7680

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 7740

gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 7800

gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttcgga 7860

ccctgcccgt taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgcttttcc 7920

gctgcataac cctgcttcgg ggtcattata gcgatttttt cggatatatcc atcctttttc 7980

gcacgatata caggattttg ccaaaggggt cgtgtagact ttccttgggtg tatccaacgg 8040

cgtcagccgg gcaggatagg tgaagtaggc caaccgcga gcgggtgttc cttcttcact 8100

gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccggct 8160

accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag 8220

ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag 8280

gcggcgggcg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca gggctacaaa 8340

atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa tggcgacctg 8400

ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac ggcgcgggttc 8460

ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc 8520

aaggtcatga tgggcgtggg ccgcccagag gcagagccat gactttttta gccgctaaaa 8580

cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga 8640

cttcgcggag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga 8700

cgctcaccgg gctggttgcc ctcgccgctg ggctggcggc cgtctatggc cctgcaaacg 8760

cgccagaaac gccgtcgaag ccgtgtgcga gacaccgagg ccgccggcgt tgtggatacc 8820

tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact tgagggggccg 8880

actcaccggg cgcggcgttg acagatgagg ggcaggctcg atttcggccg gcgacgtgga 8940

gctggccagc ctcgcaaac gccgaaaacg cctgatttta cgcgagtttc ccacagatga 9000

tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc gcgactactg 9060

acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga tgaggggcgc 9120

acctattgac atttgagggg ctgtccacag gcagaaaac cagcatttgc aagggtttcc 9180

gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca atatttataa 9240

accttgtttt taaccagggc tgcgccctgt gcgcgtgacc gcgcacgccg aaggggggtg 9300

cccccccttc tcgaaccctc ccggcccgtt aacgcgggcc tcccatcccc ccaggggctg 9360

cgcccctcgg ccgcgaacgg cctcacccca aaaatggcag cgctggcagt ccttgccatt 9420

gccgggatcg gggcagtaac gggatgggag atcagcccga gcgcgacgcc cggaagcatt 9480

gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag tgagggcggc 9540

ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga cttcatggcg 9600

gggccggcaa tttttacctt gggcattcctt ggcatagtgg tcgcgggtgc cgtgctcgtg 9660
ttcgggggtg cgataaaccc agcgaaccat ttgaggtgat aggtaagatt ataccgaggt 9720
atgaaaacga gaattggacc tttacagaat tactctatga agcgccatat ttaaaaagct 9780
accaagacga agaggatgaa gaggatgagg aggagattg ccttgaatat attgacaata 9840
ctgataagat aatatatcctt ttatatagaa gatatcgccg tatgtaagga tttcaggggg 9900
caaggcatag gcagcgcgct tatcaatata tctatagaat gggcaaagca taaaaacttg 9960
catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt ctatcataat 10020
tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgccc gatgactttg 10080
tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg tgccaggtgc 10140
tgccctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct gattacgtgc 10200
agctttccct tcaggcgga ttcatacagc ggccagccat ccgtcatcca tatcaccagc 10260
tcaaaggggtg acagcaggct cataagacgc cccagcgtcg ccatagtgcg ttcaccgaat 10320
acgtgcgcaa caaccgtctt ccggagactg tcatacgcgt aaaacagcca gcgctggcgc 10380
gatttagccc cgacatagcc cactgttcg tccatttccg cgcagacgat gacgtcactg 10440
cccggctgta tgcgcgaggt taccgactgc ggctgagtt ttttaagtga cgtaaaatcg 10500
tgttgaggcc aacgcccata atgcgggctg ttgcccgga tccaacgcca ttcattggcca 10560
tatcaatgat tttctggtgc gtaccgggtt gagaagcgtt gtaagtgaac tgcagttgcc 10620
atgttttacg gcagtgaag cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10680
gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10740
cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc cataattgtg 10800
gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac aactttgaaa 10860

aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 10920
tggtataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 10980
aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11040
aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag aaaatgaaaa 11100
cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11160
aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc tgcactttga 11220
acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc tttgctcgga 11280
agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcacag 11340
gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11400
agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11460
agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaggcc 11520
cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct ttgtgaaaga 11580
tggtcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggtatga 11640
cattgccttc tgcgtccggt cgatcaggga ggatatacggg gaagaacagt atgtcgagct 11700
atTTTTtgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 11760
ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 11820
ccgacttctt ccgcatcaag tgTTTTggct ctCaggccga ggcccacggc aagtatttgg 11880
gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac gagaaggacg 11940
gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12000
caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc ggggcaatcc 12060
cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12120

acgcgggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc atgcgtgcgc 12180
cccgcgaaac cttccagtcc gtcggctcga tgggtccagca agctacggcc aagatcgagc 12240
gcgacagcgt gcaactggct cccctgccc tgcccgcgcc atcgggccgc gtggagcgtt 12300
cgcgtcgtct cgaacaggag gcggcaggtt tggcgaagtc gatgaccatc gacacgcgag 12360
gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa caggtcagcg 12420
aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa atgcagcttt 12480
ccttgttcga tattgcgcgc tggccggaca cgatgcgagc gatgccaaac gacacggccc 12540
gctctgcctt gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg caaaacaagg 12600
tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcgag ctgcggggccg 12660
acgatgacga actggtgtgg cagcaggtgt tggagtacgc gaagcgcacc cctatcggcg 12720
agccgatcac cttcacgttc tacgagcttt gccaggacct gggctggctg atcaatggcc 12780
ggtattacac gaaggccgag gaatgcctgt cgcgccctaca ggcgacggcg atgggcttca 12840
cgtccgaccg cgttgggcac ctggaatcgg tgctgctgct gcaccgcttc cgcgtcctgg 12900
accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc gtcgtgctgt 12960
ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg tcgccgacgg 13020
cccgacggat gttcgactat ttcagctcgc accgggagcc gtacccgctc aagctggaaa 13080
ccttcgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc gagcaggtcg 13140
gcgaagcctg cgaagagttg cgaggcagcg gcctgggtgga acacgcctgg gtcaatgatg 13200
acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg ggttcagcag 13260
ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact tgcttcgctc 13320
agtatcgctc gggacgcacg gcgcgtcta cgaactgccg ataaacagag gattaaaatt 13380

gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc aggatttccg 13440
cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg tttacgagca 13500
cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg tggcattcgg 13560
cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg acggccccaa 13620
ggacgctcac aaggcgcac tgtccggcgt tttcgtggag cccgaacagc gaggccgagg 13680
ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga tgatcgtccg 13740
acagattcca acgggaatct ggtggatgcg catcttcac ctcggcgcac ttaatatctc 13800
gctattctgg agcttggtgt ttatttcggt ctaccgctg ccgggcgggg tcgcgcgac 13860
ggtaggcgt gtgcagccgc tgatggtcgt gttcatctct gccgctctgc taggtagccc 13920
gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg cgctgttggt 13980
gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg cgggggcggt 14040
ttccatggcg ttcggaaccg tgctgacccg caagtggcaa cctcccgtgc ctctgctcac 14100
ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag ctttagtggt 14160
tgatccgcca atcccgatgc ctacaggaac caatgttctc ggctggcgt ggctcggcct 14220
gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac tcgaacctac 14280
agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc cgggggatgca 14340
tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag caatggatag 14400
gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcage ttcctcagcg 14460
gctttatcca gcgatttctt attatgtcgg catagtcttc aagatcgaca gcctgtcacg 14520
gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg agatgatatt 14580

tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 14640
atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 14700
gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 14760
ctgtatcgag tggtgathtt gtgccgagct gccggtcggg gagctgttgg ctggctgggt 14820
gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac acattgcgga 14880
cgtttttaat gtactggggg ggtttttctt ttcaccagtg agacgggcaa cagctgattg 14940
cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg ttgccccagc 15000
aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat aaatcaaaag 15060
aatagcccgga gatagggttg agtggtgttc cagtttgga caagagtcca ctattaaaga 15120
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc cactacgtg 15180
aaccatcacc caaatcaagt tttttggggg cgagggtgccg taaagcacta aatcggaacc 15240
ctaaaggagg ccccgathtt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 15300
aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15360
tcgggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 15420
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtga 15480
ttcgagctcg gtaccggggg atctttcgac actgaaatac gtcgagcctg ctccgcttgg 15540
aagcggcgag gagcctcgtc ctgtcacaac taccaacatg gactacgata agggccagtt 15600
ccgccagctc attaagagcc agttcatggg cgttggcatg atggccgtca tgcattctga 15660
cttcaagtac accaacgctc ttctgatcca gtcgatcatc cgctgaaggc gctttcgaat 15720
ctggttaaga tccacgtctt cgggaagcca gcgactggtg acctccagcg tccctttaag 15780
gctgccaaca gctttctcag ccagggccag cccaagaccg acaaggcctc cctccagaac 15840

gccgagaaga actggagggg tggtgtcaag gaggagtaag ctccttattg aagtcggagg 15900
acggagcggg gtcaagagga tattcttcga ctctgtatta tagataagat gatgaggaat 15960
tgagggtagc atagcttcat ttggatttgc tttccaggct gagactctag cttggagcat 16020
agagggctct ttggctttca atattctcaa gtatctcgag tttgaactta ttccctgtga 16080
accttttatt caccaatgag cattggaatg aacatgaatc tgaggactgc aatcgccatg 16140
aggttttcga aatacatccg gatgtcgaag gcttggggca cctgcgttgg ttgaatttag 16200
aacgtggcac tattgatcat ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa 16260
cgttgctagc agttccaggg ggaatgttat gatgagcatt gtattaaatc aggagatata 16320
gcatgatctc tagttagctc accacaaaag tcagacggcg taaccaaag tcacacaaca 16380
caagctgtaa ggatttcggc acggctacgg aagacggaga agccaccttc agtggactcg 16440
agtaccattt aattctattt gtgtttgatc gagacctaata acagccccta caacgaccat 16500
caaagtcgta tagctaccag tgaggaagtg gactcaaatc gacttcagca acatctcctg 16560
gataaaacttt aagcctaaac tatacagaat aagatagggtg gagagcttat accgagctcc 16620
caaatctgtc cagatcatgg ttgaccgggtg cctggatctt cctatagaat catccttatt 16680
cgttgacctt gctgattctg gactgacctt gaggggtcatg acttgagcct aaaatccgcc 16740
gcctccacca tttgtagaaa aatgtgacga actcgtgagc tctgtacagt gaccgggtgac 16800
tctttctggc atgcggagag acggacggac gcagagagaa gggctgagta ataagccact 16860
ggccagacag ctctggcggc tctgaggtgc agtggatgat tattaatccg ggaccggccg 16920
cccctccgcc ccgaagtgga aaggctgggtg tgcccctcgt tgaccaagaa tctattgcat 16980
catcgagaa tatggagctt catcgaatca ccggcagtaa gcgaaggaga atgtgaagcc 17040
aggggtgtat agccgtcggc gaaatagcat gccattaacc taggtacaga agtccaattg 17100

cttccgatct ggtaaaagat tcacgagata gtaccttctc cgaagtaggt agagcgagta 17160
cccggcgcgt aagctcccta attggcccat ccggcatctg tagggcggtcc aaatatcgtg 17220
cctctcctgc tttgcccggg gtatgaaacc ggaaaggccg ctcaggagct ggccagcggc 17280
gcagaccggg aacacaagct ggcagtcgac ccattccggg ctctgcactc gacctgctga 17340
ggtcctcag tccctggtag gcagctttgc cccgtctgtc cgcccggtgt gtcggcgggg 17400
ttgacaaggt cgttgcgta gtccaacatt tgttgccata ttttctgct cttccaccca 17460
gctgctcttt tcttttctct ttcttttccc atcttcagta tattcatctt cccatccaag 17520
aacctttatt tcccctaagt aagtactttg ctacatccat actccatcct tccatccct 17580
tattcctttg aacctttcag ttcgagcttt ccacttcat cgcagcttga ctaacagcta 17640
ccccgcttga gcagacatca ccattgctga actcaccgag acgtctgtcg agaagtttct 17700
gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggaggggc agaatactcg 17760
tgctttcagc ttcgatgtag gagggcggtg atatgtcctg cgggtaaata gctgcgccga 17820
tggtttctac aaagatcgtt atgtttatcg gcactttgca tcggccgcgc tccgattcc 17880
ggaagtgctt gacattgggg aattcagcga gagcctgacc tattgcatct cccgccgtgc 17940
acagggtgtc acgttgcaag acctgcctga aaccgaactg cccgctgttc tgcagccggg 18000
cgcgaggagg atggatgcga tcgctgcggc cgatcttagc cagacgagcg ggttcggccc 18060
attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat gcgcgattgc 18120
tgatcccat gtgtatcact ggcaaactgt gatggacgac accgtcagtg cgtccgtcgc 18180
gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc ggcacctcgt 18240
gcacgcggat ttcggctcca acaatgtcct gacggacaat ggccgcataa cagcgggtcat 18300
tgactggagc gaggcgatgt tcggggattc ccaatacgag gtcgccaaca tcttctctg 18360

gagggcgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga ggcacccgga 18420
gcttgcagga tcgccgcggc tccgggcgta tatgctccgc attggtcttg accaactcta 18480
tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc gatgcgacgc 18540
aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca gaagcgcggc 18600
cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac gccccagcac 18660
tcgtccgagg gcaaaggaat agagtagatg ccgaccgcgg gatcgatcca cttaacgtta 18720
ctgaaatcat caaacagctt gacgaatctg gatataagat cgttgggtgc gatgtcagct 18780
ccggagttga gacaaatggg gttcaggatc tcgataagat acgttcattt gtccaagcag 18840
caaagagtgc cttctagtga tttaatagct ccatgtcaac aagaataaaa cgcgttttcg 18900
ggtttacctc ttccagatac agctcatctg caatgcatta atgcattgac tgcaacctag 18960
taacgccttn caggctccgg cgaagagaag aatagcttag cagagctatt ttcatttttcg 19020
ggagacgaga tcaagcagat caacggtcgt caagagacct acgagactga ggaatccgct 19080
cttggctcca cgcgactata ttttgtctc taattgtact ttgacatgct cctcttcttt 19140
actctgatag cttgactatg aaaattccgt caccagcncc tgggttcgca aagataattg 19200
catgtttctt ccttgaactc tcaagcctac aggacacaca ttcatcgtag gtataaacct 19260
cgaaatcant tcctactaag atggtataca atagtaacca tgcattggtt cctagtgaat 19320
gctccgtaac acccaatagc ccggccgaaa cttttttaca actctcctat gagtcgttta 19380
cccagaatgc acaggtacac ttgttttagag gtaatccttc tttctagcta gaagtcctcg 19440
tgtactgtgt aagcgccac tccacatctc cactcgacct gcaggcatgc a 19491

<210> 46

<211> 21300

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 46

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt	60
cctgtcacia ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct	240
tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600

gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660

ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720

tccgatagct ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780

tggaaatgta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacaaaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gaggaccatt taattctatt 960

tgtgtttgat cgagaccta tacagcccct acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaa at cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggg gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgaccc agagggtcat gacttgagcc taaaatccgc cgctccacc atttgtagaa 1260

aaatgtgacg aactcgtgag ctctgtacag tgaccgggtga ctctttcttg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacgggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgac tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgct caaatatcgt gcctctcctg ctttgcccg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgtc 1920

agtccaacat ttgttgccat attttcctgc tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040

taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100

gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160

accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220

gtctccgacc tgatgcagct ctccggagggc gaagaatctc gtgctttcag cttcgatgta 2280

ggagggcggtg gatatgtcct gcgggtaaatt agctgcgccg atggtttcta caaagatcgt 2340

tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgtc tgacattggg 2400

gaattcagcg agagcctgac ctattgcac tcccgccgtg cacagggtgt cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520

atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580

ggtcaataca ctacatggcg tgatttcata tgccgcgattg ctgatcccca tgtgtatcac 2640

tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700

atgctttggg ccgaggactg ccccgaaagtc cggcacctcg tgcacgcgga ttccggctcc 2760

aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820

ttcgggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880

atggagcagc agacgcgcta ctccgagcgg aggcattccg agcttgcagg atcgccgcgg 2940

ctccggggcgt atatgctccg cattgggtctt gaccaactct atcagagctt ggttgacggc 3000

aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctccg atccggagcc 3060

gggactgtcg ggcgtacaca aatcgccccg agaagcgcgg ccgtctggac cgatggctgt 3120

gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcggtccgag ggcaaaggaa 3180

tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ctttctagtg 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaacctt gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggtctc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttctactaa 3780

gatggtatac aatagtaacc atgcatggtt gcctagttaa tgctccgtaa caccaatac 3840

gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggta 3900

cttggtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgcca 3960

ctccacatct ccactcgacc tgcaggcatg caagcttgaa ttcgagctcg gtaccgggg 4020

atctttcgac actgaaatac gtcgagcctg ctccgcttgg aagcggcgag gagcctcgtc 4080

ctgtcacaac taccaacatg gactacgata agggccagtt ccgccagctc attaagagcc 4140

agttcatggg cgttggcatg atggccgtca tgcattctgta cttcaagtac accaacgctc 4200

ttctgatcca gtcgatcatc cgctgaaggc gctttcgaat ctgggtaaga tccacgtctt 4260

cggaagcca gcgactgggtg acctccagcg tccctttaag gctgccaaca gctttctcag 4320

ccagggccag cccaagaccg acaaggcctc cctccagaac gccgagaaga actggagggg 4380

tggtgtcaag gaggagtaag ctcttattg aagtcggagg acggagcggg gtcaagagga 4440

tattcttcga ctctgtatta tagataagat gatgaggaat tggaggtagc atagcttcat 4500

ttggatttgc tttccaggct gagactctag cttggagcat agagggtcct ttggctttca 4560

atattctcaa gtatctcgag tttgaactta ttccctgtga accttttatt caccaatgag 4620

cattggaatg aacatgaatc tgaggactgc aatcgccatg aggttttcga aatacatccg 4680

gatgtcgaag gcttggggca cctgcgttgg ttgaatttag aacgtggcac tattgatcat 4740

ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa cgttgctagc agttccagg 4800

ggaatgttat gatgagcatt gtattaaatc aggagatata gcatgatctc tagttagctc 4860

accacaaaag tcagacggcg taacacaaaag tcacacaaca caagctgtaa ggatttcggc 4920

acggctacgg aagacggaga agccaccttc agtggactcg agtaccattt aattctat 4980

gtgtttgatc gagacctaata acagccccta caacgaccat caaagtcgta tagctaccag 5040

tgaggaagtg gactcaaata gacttcagca acatctctg gataaacttt aagcctaaac 5100

tatacagaat aagatagggtg gagagcttat accgagctcc caaatctgtc cagatcatgg 5160

ttgaccgggtg cctggatctt cctatagaat catccttatt cgttgaccta gctgattctg 5220

gagtgaccca gagggatcatg acttgagcct aaaatccgcc gcctccacca tttgtagaaa 5280

aatgtgacga actcgtgagc tctgtacagt gaccgggtgac tctttctggc atgcggagag 5340

acggacggac gcagagagaa gggctgagta ataagccact ggccagacag ctctggcggc 5400

tctgagggtg agtggatgat tattaatccg ggaccggccg cccctccgcc ccgaagtgga 5460

aaggctggtg tgcccctcgt tgaccaagaa tctattgcat catcggagaa tatggagctt 5520

catcgaatca ccggcagtaa gcgaaggaga atgtgaagcc aggggtgtat agccgtcggc 5580

gaaatagcat gccattaacc taggtacaga agtccaattg cttccgatct ggtaaaagat 5640
tcacgagata gtaccttctc cgaagtaggt agagcgagta cccggcgcggt aagctcccta 5700
attggcccat cggcatctg tagggcggtcc aaatatcgtg cctctcctgc tttgcccgggt 5760
gtatgaaacc ggaaaggccg ctccaggagct ggccagcggc gcagaccggg aacacaagct 5820
ggcagtcgac ccatccgggtg ctctgcactc gacctgctga ggtccctcag tccctggtag 5880
gcagctttgc cccgtctgtc cgcccgggtg gtcggcgggg ttgacaaggt cgttgcgtca 5940
gtccaacatt tgttgccata ttttcctgct ctccccacca gctgctcttt tcttttctct 6000
ttcttttccc atcttcagta tattcatctt cccatccaag aacctttatt tcccctaagt 6060
aagtactttg ctacatccat actccatcct tcccatccct tattcctttg aacctttcag 6120
ttcgagcttt ccacttcat cgcagcttga ctaacagcta ccccgcttga gcagacatca 6180
ccatgtcaat actcacttat ctggaatttc atctctacta tacactacct gtccttgcggt 6240
cattgtgttg gctgctaaag ccgtttcact cacagcaaga caatctcaag tataaatttt 6300
taatgttgat ggccgcctct accgcatcga tttgggacaa ttatatcggt tatcatcgcg 6360
cttggtggtg ctgtcctact tgtgttggtg ctgtcattgg ctatgtacct ctagaagaat 6420
acatgttctt tatcatcatg actttaatga ctgtcgcgtt ctcaaacttt gttatgcgtt 6480
ggcacttgca tactttcttt attagacca acacttcttg gaagcaaaca ctattagtac 6540
gccttgtgcc tgtttcagct ttattggcaa tcacttatca tgcttggcac ttgacactgc 6600
caaataaacc ttcattttat ggttcatgca tcctttggta tgcttgtcct gtgttggcta 6660
ttctttgggt ggggtgctggc gaatatatct tgcgtcgacc tgtggctgtc cttttgtcta 6720
ttgttatccc tagtgtatac ctatgttggg ctgatatcgt cgctattagt gctggcacat 6780
ggcatatttc tcttagaaca agcactggca aaatggtagt acccgattta cctgtagaag 6840

aatgcctggt ttttactttg atcaacacag tcttggtttt tgctacctgt gctatagacc 6900
gcgctcaggc catcctccat gtgagcgcgc gtaatacgac tcactatagg gcgaattgga 6960
gctccaccgc ggtggcggcc gctctagaac tagtggatcc cccgggctgc aggaattcgg 7020
cacgagctac atttcacaag cccgtgagcg gtgcaagcgc tctgccccac atcggccccac 7080
ctcctcatct ccatcgggtca tttgctgcta ccacgatgct gtcgaagctg cagtcaatca 7140
gcgtcaaggc ccgccgcgtt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc 7200
atgctcagcg gtgctcgta gtccggctgc gactggcagc accacagaca gaggaggcgc 7260
tggaaccgt gcaggctgcc ggcgcgggcg atgagcacag cgccgatgta gcactccagc 7320
agcttgaccg ggctatcgca gagcgtcgtg cccggcgcaa acgggagcag ctgtcatacc 7380
aggctgccgc cattgcagca tcaattggcg tgtcaggcat tgccatcttc gccacctacc 7440
tgagatttgc catgcacatg accgtgggcg gcgcagtgcc atggggtgaa gtggctggca 7500
ctctcctctt ggtggttggg ggcgcgctcg gcatggagat gtatgccgc tatgcacaca 7560
aagccatctg gcatgagtcg cctctgggct ggctgctgca caagagccac cacacacctc 7620
gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc 7680
tcctgtgtac ctttggcttc tggctgcca acgtcctggg ggcggcctgc tttggagcgg. 7740
ggctgggcat cacgtatac ggcatggcat atatgtttgt acacgatggc ctggtgcaca 7800
ggcgctttcc caccgggccc atcgctggcc tgccctacat gaagcgccctg acagtggccc 7860
accagctaca ccacagcggc aagtacgggt gcgcgccctg gggatatgttc ttgggtccac 7920
aggagctgca gcacattcca ggtgcggcgg aggaggtgga gcgactggtc ctggaactgg 7980
actggtccaa gcgggctcag gccatcctcc atctgtacaa atcatctgtt caaaatcaaa 8040
accctaaaca agccatttcc cttttccagc atgtcaaaga gctagcatgg gccttctgtc 8100

ttcctgacca aatgctcaac aatgaattgt ttgatgatct tactatcage tgggatatatt 8160

tacgtaaagc ctcaaagtca ttctatactg catctgccgt ttttccaagt tatgtacgtc 8220

aagacttggg tgttctctat gctttctgca gagctaccga tgacctgtgc gatgatgaat 8280

ccaaatctgt tcaagaaaga agagaccaat tagatcttac tcgacaattt gttcgtgatc 8340

tctttagcca aaagaccagt gcgcctattg tgattgattg ggaattgtat caaaaccaac 8400

ttcctgcttc ttgtatatca gccttttagag cctttactcg ccttcgccat gtccttgaag 8460

tagaccctgt agaagaacta ttagatgggt acaaatggga tcttgagcgt cgtcctatcc 8520

ttgatgaaca agacttggag gcatactctg cttgtgtggc cagtagtgtg ggtgaaatgt 8580

gcacacgtgt gattcttgct caagaccaa aggaaaatga tgcttgata attgaccgtg 8640

cacgtgagat ggggctggtg ctacaatacg ttaacattgc tcgagacatt gtgactgata 8700

gcgagactct gggtcgatgt tatctgcctc aacaatggct tagaaaagaa gaaacagaac 8760

aaatacagca aggcaacgcc cgtagcctag gtgatcaaag actgttgggc ttgtctctga 8820

agcttgtagg aaaggcagac gctatcatgg tgagagctaa gaagggcatt gacaagttgc 8880

cggcaaaactg tcaaggcggg gtacgagctg cttgccaaagt atatgctgca attggatctg 8940

tactcaagca gcagaagaca acatatacta caagagctca tctaaaagga agcgaacgtg 9000

ccaagattgc tctgttgagt gtatacaacc tctatcaatc tgaagacaag cctgtggctc 9060

tccgtcaagc tagaaagatt aagagttttt ttgttgatta gtgaattttt gttttattta 9120

tgtctgatag ttcaataaag agacaacaca tacaataaa aatcattgtc tttaaatggt 9180

aatttagtag agtgtaaagc ctgcattttt tttgtacgca taaacaatga gttcaccocg 9240

cttctgggtt ttaaataatt atgtcaaact agggaaaatt ctttttttct tcttcgttct 9300

ttttttggct tgttgtggag tcacaggctt gtcttcagat tgatagaggt tgtatacact 9360

caacagagca atcttggcac gttcgcttcc ttttagatga gctcttgtag gatatgttgt 9420
cttctgctgc ttgagtacag atccaattgc agcatatact tggcaagcag ctcgtacacc 9480
gccttgacag tttgccggca acttgtcaat gcccttctta gctctcacca tgatagcgtc 9540
tgccttttct acaagcttgg cgtaatcatg gtcatactgt tttcctgtgt gaaattgtta 9600
tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaa cctgggggtgc 9660
ctaagtgtg agctaactca cattaattgc gttgcgctca ctgcccgtt tccagtcggg 9720
aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg 9780
tattgggcca aagacaaaag ggcgacattc aaccgattga gggagggag gtaaatattg 9840
acggaaatta ttcattaaag gtgaattatc accgtcaccg acttgagcca tttgggaatt 9900
agagccagca aaatcaccag tagcaccatt accattagca aggccggaaa cgtcaccaat 9960
gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc ctttagcgtc 10020
agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag cgtttgccat 10080
cttttcataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc cctcagagcc 10140
gccaccctca gaaccgccac cctcagagcc accaccctca gagccgccac cagaaccacc 10200
accagagccg ccgccagcat tgacaggagg cccgatctag taacatagat gacaccgcgc 10260
gcgataattt atcctagttt gcgcgctata ttttgtttct tatcgctat taaatgtata 10320
attgcgggac tctaatacata aaaaccatc tcataaataa cgtcatgcat tacatgttaa 10380
ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca agaccggcaa 10440
caggattcaa tcttaagaaa ctttattgcc aaatgtttga acgatcgggg atcatccggg 10500
tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcgcg tgcatctcgg 10560
tcttgcttgg gcagtcgccg ccgacgccgt tgatgtggac gccgggccc atcatattgt 10620

cgctcaggat cgtggcggtg tgcttgctcg cgttgctgt cgtaatgata tcggcacctt 10680

cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga tccccgcgct 10740

ggaggatcat ccagccggcg tcccggaaaa cgattccgaa gcccacactt tcatagaagg 10800

cgggcgtgga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg gtcatttcga 10860

accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga tgcgctgcga 10920

atcgggagcg gcgataccgt aaagcacgag gaagcgggtca gccatttcgc cgccaagctc 10980

ttcagcaata tcacgggtag ccaacgctat gtctgatag cggtcgccca caccagccg 11040

gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg gcaagcaggc 11100

atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cgcgcttga gcctggcgaa 11160

cagttcggct ggcgcgagcc cctgatgctc ttcgctcaga tcctctgat cgacaagacc 11220

ggcttccatc cgagtacgtg ctcgctcgat gcgatgtttc gcttggtggg cgaatgggca 11280

ggtagccgga tcaagcgtat gcagccgccg cattgcatca gccatgatgg atactttctc 11340

ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgcca atagcagcca 11400

gtcccttccc gcttcagtga caacgtcgag cacagctcg caaggaacgc cgtcgtggc 11460

cagccacgat agccgcgctg cctcgtcctg cagttcatc agggcaccgg acaggtcgg 11520

cttgacaaaa agaaccgggc gcccctgcgc tgacagccgg aacacggcgg catcagagca 11580

gccgattgtc tggttggtccc agtcatagcc gaatagctc tccaccaag cgcccgga 11640

acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg tgcagattat 11700

ttggattgag agtgaatatg agactctaag tggataccga ggggaattta tggaacgtca 11760

gtggagcatt ttgacaaga aatatttgct agctgatagt gaccttaggc gacttttgaa 11820

cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa acccgcggt 11880
gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc ttgtcccgcg 11940
tcatcggcgg gggtcataac gtgactccct taattctccg ctcatgatca gattgtcggt 12000
tccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt aaacctaaga 12060
gaaaagagcg tttattagaa taatcggata tttaaaggc cgtgaaaagg tttatccgtt 12120
cgtccatttg tatgtgcatg ccaaccacag gggtccccag atctggcgcc ggccagcgag 12180
acgagcaaga ttggccgccc cccgaaacga tccgacagcg cgcccagcac aggtgcgcag 12240
gcaaattgca ccaacgcata cagcgccagc agaatgccat agtgggcgggt gacgtcggtc 12300
gagtgaacca gatcgcgagc gaggcccgcc agcaccggca taatcaggcc gatgccgaca 12360
gcgtcgagcg cgacagtgtc cagaattacg atcaggggta tgttgggttt cacgtctggc 12420
ctccggacca gcctccgtg gtccgattga acgcgcggat tctttatcac tgataagttg 12480
gtggacatat tatgtttatc agtgataaag tgtcaagcat gacaaagttg cagccgaata 12540
cagtgatccg tgccgccttg gacctgttga acgaggtcgg cgtagacgggt ctgacgacac 12600
gcaaactggc ggaacggttg ggggttcagc agccggcgct ttactggcac ttcaggaaca 12660
agcgggcgct gctcgacgca ctggccgaag ccatgctggc ggagaatcat acgcattcgg 12720
tgccgagagc cgacgacgac tggcgctcat ttctgatcgg gaatgccgcg agcttcaggc 12780
aggcgctgct cgcctaccgc gatggcgcgc gcatccatgc cggcacgcga ccgggcgcac 12840
cgcatatgga aacggccgac gcgcagcttc gcttctcttg cgaggcgggt ttttcggccg 12900
gggacgccgt caatgcgctg atgacaatca gctacttcac tgttggggcc gtgcttgagg 12960
agcaggccgg cgacagcgat gccggcgagc gcggcgccac cgttgaacag gctccgctct 13020
cgccgctggt gcgggcccgc atagacgcct tcgacgaagc cgggtccggac gcagcgttcg 13080

agcagggact cgcggtgatt gtcgatggat tggcgaaaag gaggctcggt gtcaggaacg 13140

ttgaaggacc gagaaagggg gacgattgat caggaccgct gccggagcgc aaccactca 13200

ctacagcaga gccatgtaga caacatcccc tcccccttc caccgcgtca gacgcccgtg 13260

gcagcccgtc acgggctttt tcatgccctg ccctagcgct caagcctcac ggccgcgctc 13320

ggcctctctg ggggccttct ggcgctcttc cgcttcctcg ctactgact cgctgcgctc 13380

ggtcgttcgg ctgcggcgag cggatatcagc tactcaaag gcggtatac gggtatccac 13440

agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 13500

ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 13560

caaaaatcga cgctcaagtc agagggtggcg aaacccgaca ggactataaa gataccaggc 13620

gtttccccct ggaagctccc tcgtgcgctc tcctgttcgg accctgcgcg ttaccggata 13680

cctgtccgcc tttctccctt cgggaagcgt ggcgcttttc cgctgcataa ccctgcttcg 13740

gggtcattat agcgattttt tcggtatatc catccttttt cgacgatat acaggatttt 13800

gccaaagggg tcgtgtagac tttccttggg gtatccaacg gcgtcagccg ggcaggatag 13860

gtgaagtagg cccaccgcg agcgggtgtt cttcttcac tgtcccttat tcgcacctgg 13920

cggtgctcaa cgggaatcct gctctgcgag gctggccggc taccgccggc gtaacagatg 13980

agggcaagcg gatggctgat gaaaccaagc caaccaggaa gggcagccca cctatcaagg 14040

tgtactgcct tccagacgaa cgaagagcga ttgaggaaaa ggccggggcg gccggcatga 14100

gcctgtcggc ctacctgctg gccgtcggcc agggctacaa aatcacgggc gtcgtggact 14160

atgagcacgt ccgcgagctg gcccgcatca atggcgacct gggccgcctg ggcggcctgc 14220

tgaaactctg gctcaccgac gaccgcgca cggcgcggtt cgggtgatgcc acgatcctcg 14280

ccctgctggc gaagatcgaa gagaagcagg acgagcttgg caaggatcatg atgggcgtgg 14340

tccgccccgag ggcagagcca tgactttttt agccgctaaa acggccgggg ggtgcgcgtg 14400
attgccaagc acgtcccat gcgctccatc aagaagagcg acttcgcgga gctggtgaag 14460
tacatcacccg acgagcaagg caagaccgag cgcctttgcg acgctcacccg ggctgggttc 14520
cctcgccgct gggctggcgg ccgtctatgg ccctgcaaac gcgccagaaa cgccgtcgaa 14580
gccgtgtgcg agacaccgcg gccgccggcg ttgtggatac ctgcgggaaa acttggccct 14640
cactgacaga tgagggggcg acgttgacac ttgaggggccc gactcacccg gcgcggcgctt 14700
gacagatgag gggcaggctc gatttcggcc ggcgacgtgg agctggccag cctcgcaaat 14760
cggcgaaaac gcctgatttt acgcgagttt cccacagatg atgtggacaa gcctggggat 14820
aagtgccctg cggatttgac acttgagggg cgcgactact gacagatgag gggcgcgatc 14880
cttgacactt gaggggcaga gtgctgacag atgaggggcg cacctattga catttgaggg 14940
gctgtccaca ggcagaaaat ccagcatttg caagggttcc cgcccgtttt tcggccaccg 15000
ctaacctgtc ttttaacctg cttttaaaacc aatatttata aaccttgttt ttaaccaggg 15060
ctgcgccctg tgcgcgtgac cgcgcacgcc gaaggggggt gccccccctt ctgaacct 15120
cccggcccg taacgcgggc ctcccatccc ccagggggct gcgcccctcg gccgcgaacg 15180
gcctcacccc aaaaatggca gcgctggcag tccttgccat tgccgggatc ggggcagtaa 15240
cgggatgggc gatcagcccg agcgcgacgc ccggaagcat tgacgtgccg caggtgctgg 15300
catcgacatt cagcgaccag gtgccgggca gtgagggcgj cggcctgggt ggcggcctgc 15360
ccttcacttc ggccgtcggg gcattcacgg acttcatggc ggggccggca atttttacct 15420
tgggcattct tggcatagt gtcgcggtg ccgtgctcgt gtgcgggggt gcgataaacc 15480
cagcgaacca tttaggtga taggtaagat tataccgagg tatgaaaacg agaattggac 15540
ctttacagaa ttactctatg aagcgccata tttaaaaagc taccaagacg aagaggatga 15600

agaggatgag gaggcagatt gccttgaata tattgacaat actgataaga taatatatct 15660
tttatataga agatatcgcc gtatgtaagg atttcagggg gcaaggcata ggcagcgcgc 15720
ttatcaatat atctatagaa tgggcaaagc ataaaaactt gcatggacta atgcttgaaa 15780
cccaggacaa taaccttata gcttgtaaatt tctatcataa ttgggtaatg actccaactt 15840
attgatagtg ttttatgttc agataatgcc cgatgacttt gtcatgcagc tccaccgatt 15900
ttgagaacga cagcgacttc cgtcccagcc gtgccagggt ctgcctcaga ttcagggttat 15960
gccgctcaat tcgctgcgta tctcgcttgc tgattacgtg cagctttccc ttcaggcggg 16020
attcatacag cggccagcca tccgtcatcc atatcaccac gtcaaagggt gacagcaggc 16080
tcataagacg ccccagcgtc gccatagtgc gttcaccgaa tacgtgcgca acaaccgtct 16140
tccggagact gtcatacgcg taaaacagcc agcgtggcg cgatttagcc ccgacatagc 16200
cccactgttc gtccatttcc ggcagacga tgacgtcact gcccggtgt atgcgcgagg 16260
ttaccgactg cggcctgagt tttttaagt acgtaaaatc gtgttgaggc caacgccccat 16320
aatgcgggct gttgcccgcc atccaacgcc attcatggcc atatcaatga ttttctggtg 16380
cgtaccgggt tgagaagcgg tgtaagtga ctgcagttgc catgttttac ggcagtgaga 16440
gcagagatag cgctgatgtc cggcgtgtct tttgccgta cgcaccaccc cgtcagtagc 16500
tgaacaggag ggacagctga tagacacaga agccactgga gcacctcaa aacaccatca 16560
tacactaaat cagtaagttg gcagcatcac ccataattgt ggtttcaaaa tcggctccgt 16620
cgatactatg ttatacgcca actttgaaaa caactttgaa aaagctgttt tctggtattt 16680
aaggtttttag aatgcaagga acagtgaatt ggagttcgtc ttgttataat tagcttcttg 16740
gggtatcttt aaatactgta gaaaagagga aggaaataat aaatggctaa aatgagaata 16800
tcaccggaat tgaaaaaact gatcgaaaaa taccgctgcg taaaagatac ggaaggaaatg 16860

tctcctgcta aggtatataa gctggtggga gaaaatgaaa acctatattt aaaaatgacg 16920
gacagccggt ataaaggac cacctatgat gtggaacggg aaaaggacat gatgctatgg 16980
ctggaaggaa agctgcctgt tccaaaggtc ctgcactttg aacggcatga tggctggagc 17040
aatctgctca tgagtgaggc cgatggcgtc ctttgctcgg aagagtatga agatgaacaa 17100
agccctgaaa agattatcga gctgtatgcg gagtgcacat ggctctttca ctccatcgac 17160
atatcggatt gtccctatac gaatagctta gacagccgct tagccgaatt ggattactta 17220
ctgaataacg atctggccga tgtggattgc gaaaactggg aagaagacac tccatttaaa 17280
gatccgcgcg agctgtatga ttttttaaag acggaaaagc ccgaagagga acttgtcttt 17340
tcccacggcg acctgggaga cagcaacatc tttgtgaaag atggcaaagt aagtggcttt 17400
attgatcttg ggagaagcgg cagggcggac aagtggatg acattgcctt ctgcgtccgg 17460
tcgatcaggg aggatatcgg ggaagaacag tatgtcgagc tattttttga ctactgggg 17520
atcaagcctg attgggagaa aataaaatat tatattttac tggatgaatt gtttttagtac 17580
ctagatgtgg cgcaacgatg ccggcgacaa gcaggagcgc accgacttct tccgcatcaa 17640
gtgttttggt tctcaggccg aggccacgg caagtatttg ggcaaggggt cgctggtatt 17700
cgtgcagggc aagattcgga ataccaagta cgagaaggac ggccagacgg tctacgggac 17760
cgacttcatt gccgataagg tggattatct ggacaccaag gcaccaggcg ggtcaaatca 17820
ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag ggtgaatgaa 17880
tcggacgttt gaccggaagg catacaggca agaactgac gacgcggggg tttccgccga 17940
ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg ccccgcgaaa ccttccagtc 18000
cgtcggctcg atggtccagc aagctacggc caagatcgag cgcgacagcg tgcaactggc 18060
tccccctgcc ctgcccgcg catcggccgc cgtggagcgt tcgcgtcgtc tcgaacagga 18120

ggcggcaggt ttggcgaagt cgatgaccat cgacacgcga ggaactatga cgaccaagaa 18180
gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc gaggccaagc aggccgcgtt 18240
gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttgttcg atattgcgcc 18300
gtggccggac acgatgagc cgatgccaaa cgacacggcc cgctctgcc tgttcaccac 18360
gcgcaacaag aaaatcccgc gcgaggcgct gcaaaacaag gtcattttcc acgtcaacaa 18420
ggacgtgaag atcacctaca ccggcgtcga gctgcgggcc gacgatgacg aactgggtgtg 18480
gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca ccttcacgtt 18540
ctacgagctt tgccaggacc tgggctggtc gatcaatggc cgggtattaca cgaaggccga 18600
ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc acgtccgacc gcgttgggca 18660
cctggaatcg gtgtcgctgc tgcaccgctt ccgcgtcctg gaccgtggca agaaaacgtc 18720
ccgttgccag gtcctgatcg acgaggaaat cgtcgtgctg tttgctggcg accactacac 18780
gaaattcata tgggagaagt accgcaagct gtcgccgacg gcccgacgga tgttcgacta 18840
tttcagctcg caccgggagc cgtaccgct caagctggaa accttccgcc tcattgtcgg 18900
atcggattcc accgcgtga agaagtggcg cgagcaggtc ggcgaagcct gcgaagagtt 18960
gcgaggcagc ggcttgggtg aacacgcctg ggtcaatgat gacctggtgc attgcaaacg 19020
ctagggcctt gtggggtcag ttccggctgg gggttcagca gccagcgctt tactggcatt 19080
tcaggaacaa gcgggcactg ctcgacgcac ttgcttcgct cagtatcgct cgggacgcac 19140
ggcgcgctct acgaactgcc gataaacaga ggattaaaat tgacaattgt gattaaggct 19200
cagattcgac ggcttggagc ggccgacgtg caggatttcc gcgagatccg attgtcggcc 19260
ctgaagaaag ctccagagat gttcgggtcc gtttacgagc acgaggagaa aaagcccatg 19320

gagggcgttcg ctgaacgggtt gcgagatgcc gtggcattcg gcgcctacat cgacggcgag 19380
atcattgggc tgtcgggtctt caaacaggag gacggcccca aggacgctca caaggcgcat 19440
ctgtccggcg ttttcgtgga gcccgaaacag cgaggccgag gggtcgccgg tatgtctgctg 19500
cgggcggttgc cggcggggtt attgctcgtg atgatactcc gacagattcc aacgggaatc 19560
tgggtggatgc gcatcttcat cctcggcgca cttaatatct cgctattctg gagcttggtg 19620
tttatttcgg tctaccgcct gccggggcgg gtcgcggcga cggtaggcgc tgtgcagccg 19680
ctgatggtcg tgttcatctc tgcgcctctg ctaggtagcc cgatacgatt gatggcggtc 19740
ctggggggcta tttgcggaac tgcgggcgtg gcgctggttg tgttgacacc aaacgcagcg 19800
ctagatcctg tcggcgtcgc agcgggcctg gcgggggcgg tttccatggc gttcggaacc 19860
gtgctgaccc gcaagtggca acctcccgtg cctctgctca cctttaccgc ctggcaactg 19920
gcggccggag gacttctgct cgttccagta gctttagtgt ttgatccgcc aatcccgatg 19980
cctacaggaa ccaatgttct cggcctggcg tggctcggcc tgatcggagc gggtttaacc 20040
tacttccttt ggttccgggg gatctcgcga ctcgaacctc cagttgtttc cttactgggc 20100
tttctcagcc ccagatctgg ggtcgatcag ccgggggatgc atcaggccga cagtcggaac 20160
ttcgggtccc cgacctgtac cattcgggtga gcaatggata ggggagttga tatcgtcaac 20220
gttcacttct aaagaaatag cgccactcag cttcctcagc ggctttatcc agcgatttcc 20280
tattatgtcg gcatagttct caagatcgac agcctgtcac ggttaagcga gaaatgaata 20340
agaaggctga taattcggat ctctgcgagg gagatgatat ttgatcacag gcagcaacgc 20400
tctgtcatcg ttacaatcaa catgctaccc tccgcgagat catccgtgtt tcaaaccggg 20460
cagcttagtt gccgttcttc cgaatagcat cggtaacatg agcaaagtct gccgccttac 20520
aacggctctc ccgctgacgc cgtcccggac tgatgggctg cctgtatcga gtgggtgattt 20580

tgtgccgagc tgccggtcgg ggagctgttg gctggctggt ggcaggatat atttgtggtgt 20640
aaacaaattg acgcttagac aacttaataa cacattgcgg acgtttttaaa tgtactgggg 20700
tggtttttct tttcaccagt gagacgggca acagctgatt gcccttcacc gcctggccct 20760
gagagagttg cagcaagcgg tccacgctgg ttgccccag caggcgaaaa tcctgtttga 20820
tggtggttcc gaaatcggca aaatccctta taaatcaaaa gaatagcccg agatagggtt 20880
gagtgttggt ccagtttggg acaagagtcc actattaaag aacgtggact ccaacgtcaa 20940
agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac ccaaatcaag 21000
ttttttgggg tcgaggtgcc gttaaagcact aaatcggaac cctaaaggga gccccgatt 21060
tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga aagcgaaagg 21120
agcgggcgcc attcaggctg cgcaactgtt gggaaggcg atcgggtgcgg gcctcttcgc 21180
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag 21240
ggttttccca gtcacgacgt tgtaaaacga cggccagtga attcgagctc ggtacccggg 21300

<210> 47

<211> 17756

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 47

cggggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gggccgccg gcgttggtga tacctcgccg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
tttctggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc	660
tcggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgca cggccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg	840
ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg	960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttataaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatacttata gaatgggcaa agcataaaaaa cttgcatgga 1260

ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtccgcatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatalc gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattcaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtgggtttca 2100

aatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta tttaaggttt tagaatgcaa ggaacagtga attggagtgc gtcttgttat 2220

aattagcttc ttggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640
tcaactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgct ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggctgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggctgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360
gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccc cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgtctctg 3840

ccctgttcac caccgcgaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtggtggagt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg ctgtttctgtg 4260

gcgaccacta cagaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgcg 4320

ggatgttcga ctatttcagc tcgcacccgg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatcgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg caccggcgcg tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggctg tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctgggtga tgcgcatctt catcctcggc gcacttaata ttctgctatt 5100

ctggagcttg ttgtttattt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatttgcg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgcttgga caatggcgagg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttgggttcg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcacaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgca gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000
tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060
cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180
taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240
accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc cagcaggcga 6300
aaatcctggt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480
cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaagggg 6600
agaaaagcga aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660
cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780
ctcggtaacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840
cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900
gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960
gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020
aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgctccct taaggctgcc 7080
aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140
aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260
tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320
tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380
tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440
tcgaaataca tccggatgac gaaggcttgg ggcacctgag ttggttgaat ttagaacgtg 7500
gcactattga tcatccgata gctctgcaaa gggcggttgca caatgcaagt caaacgttgc 7560
tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620
tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680
gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740
atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800
cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860
ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920
tgtccagatc atgggtgacc ggtgcctgga tcttcctata gaatcatcct tattegttga 7980
cctagctgat tctggagtga cccagagggt catgacttga gcctaaaate cgccgcctcc 8040
accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100
tgccatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160
acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220
cgccccgaag tggaagggt ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280
agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640

tcagtccctg gtaggcagct ttgccccgtc tgtccgcccgtgtgtgcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttccct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccgcg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtcttt 9060

cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgt gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgtg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgcgccaggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggtc tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct cgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg eggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc a gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcggttg tgatcgatgc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacggtc atttgtcaa gcagcaaaga 10140
gtgccttcta gtgatttaac agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct cgggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt cattttgctt 10800
tgtaaatttc tggttaactgc caccaagaaa tatgaggata ttcgtgatgt tcctcgtggt 10860
agccaaaatg atagcacgtg ataaatgacc accaaatagg acggctaatt gtttgggcac 10920
aatgaggctg aacataaccc cctattgggt cactatgggg taaaaaagta ccaaaataga 10980

ataattgtaa tgaacttaaa agcgagggta gcacccaaaa gtaagttaga ttatcacttg 11040
ggatatggag tatgtattta gcaaagttat aaataatagt caacgcaatt atttgcccc 11100
aactccagta acctttcata aaatgaaaat accaagcaaa gaaacttttg tgtttaccat 11160
tgtgaaaatc cgggtctatt gagcttgctg gattgtgggtg gtgtaaccaa tgttttttca 11220
atagtttttg atatggtaaa agaccataaa gggatagggt caatgttcca atcaaagat 11280
taatcttggt gttttgggga aatactacgc catgcatggc atcatgagat gtaataaata 11340
atcccgata taaaaatgtt tgccatagta taacaggcaa taacatcaa aattttagct 11400
ttgagatgtc aagggaaagt aataaactca ggctaatagac ccatgcgcta acaatgacaa 11460
tagcaatgaa aagccccta aactgagatt tacttctcag tactggagtc agttttgctt 11520
gatgactgag tggttgttct aactggatca tttctaaaga gaaggtggaa caatgttagc 11580
ataattgtgc ttgagtgagg actttgaggg taggtacata cttgataaag ttaatgatta 11640
aagagaaaaa aaaagttttg gttcaaagca gaaattgttt tttaaatacga ttggtgagaa 11700
aatttttttc tgtttccgca tcaccaaagc cacctcagga atggtcacaa attattggtc 11760
tgattggacc ataagcatat aaaaagtcca ttgaagtata cttagtggct tattagactt 11820
ttatcgTTTT ctaacgcgaa tcagcaatgt ttcttgTTTT atttactgct tgcttttagat 11880
catttttgtc tgaaatatta tgcatttggt caaagcggcc tttgtttcct ttctttcatg 11940
cttaaacacg ttgtttattc catatattac tttgaatatg catcaccgca aagcggaagt 12000
gcaaaataac aaagaacctc tttgggttac acgatcaact gctattgtga aaaaaatttc 12060
tttttgaaaa tttttggaat aatatctctt gcaaaaaaga aattttgtat atttagtagc 12120
atcaagaaca aatgaaagaa gtgtgggata acaagaatac atcatcttta gacaaaagta 12180

cgagaaaaat ctaataagtt gttatagagg tctttgtttt ctttgtgttt atagacagtt 12240
athtagagtt tgaaaagtgt ctctaagtgt tcttttttta ttattattat ttcaaagtgt 12300
atgtaatata gctaaagcta tagatttgac attttttcta aatataaaat ttcagtcaac 12360
agaaataaat gacacgagtt ctttttctct ctctcaatcc tgttgatcat caatctttga 12420
tgtcgtttta aaacaaatga atggcattta gttccttagg tgtcactcac atcttgttga 12480
ccagaaaatc cttattcgcc ctcaaactctg ctttattcct ttcatttgat ttgatgttta 12540
agtaatgcaa gcaaacaaaa aagaaacctt tcttgcaaag acaaaagaat tgttttcaga 12600
ggaaagcaac tcgttgtcat tttttaagga tttagactta taatcgacac catagtttgt 12660
ccgttacatt ttttattgtc gttttctgat ttccttttaa tctttaagca aaatcaatat 12720
taacttatct tgtcttccaa taaaaaatgg ataccaataa caataaatcc ttcacaaaga 12780
aaaaaaaaaa aaactcgaaa aaagcttggc gtaatcatgg tcatagctgt ttctgtgtgt 12840
aaattgttat ccgctcaciaa ttccacacia catacgagcc ggaagcataa agtgtaaagc 12900
ctgggggtgcc taatgagtga gctaactcac attaatgctg ttgcgctcac tgcccgtttt 12960
ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgct cggggagagg 13020
cggtttgctg attgggcaa agacaaaagg gcgacattca accgattgag ggaggggaagg 13080
taaataattga cggaaattat tcattaaagg tgaattatca ccgtcaccga cttgagccat 13140
ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa ggccggaaac 13200
gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat caagtttgcc 13260
tttagcgtca gactgtagct cgttttcatc ggcattttcg gtcatagccc ccttattagc 13320
gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg aaccgcctcc 13380
ctcagagccg ccaccctcag aaccgccacc ctcagagcca ccaccctcag agccgccacc 13440

agaaccacca ccagagccgc cgccagcatt gacaggaggg ccgatctagt aacatagatg 13500

acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct atcgcgtatt 13560

aatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac gtcatgcatt 13620

acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata atcatcgcaa 13680

gaccggcaac aggattcaat cttaagaaac tttattgcc aatgtttgaa cgatcgggga 13740

tcattccgggt ctgtggcggg aactccacga aaatatccga acgcagcaag atatcgcggt 13800

gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt gatgtggacg cggggcccga 13860

tcattattgtc gctcaggatc gtggcgttgt gcttgtcggc cgttgctgtc gtaatgatat 13920

cggcaccttc gaccgcctgt tccgcagaga tcccgtgggc gaagaactcc agcatgagat 13980

ccccgcgctg gaggatcatc cagccggcgt cccggaaaac gattccgaag cccaaccttt 14040

catagaaggc ggcggtggaa tcgaaatctc gtgatggcag gttgggcgtc gcttggtcgg 14100

tcatttcgaa ccccagagtc ccgctcagaa gaactcgtca agaaggcgat agaaggcgat 14160

gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc 14220

gccaagctct tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac 14280

accagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg 14340

caagcaggca tcgccatggg tcacgacgag atcatcgccg tcgggcatgc ggccttgag 14400

cctggcgaac agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgac 14460

gacaagaccg gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttgggtggc 14520

gaatgggcag gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga 14580

tactttctcg gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa 14640

tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc 14700

cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc agttcattca gggcaccgga 14760
caggtcggtc ttgacaaaaa gaaccgggcg ccctgcgct gacagccgga acacggcggc 14820
atcagagcag ccgattgtct gttgtgcca gtcatagccg aatagcctct ccaccaagc 14880
ggccggagaa cctgcgtgca atccatcttg ttcaatcatg cgaaacgata cagatccggt 14940
gcagattatt tggattgaga gtgaatatga gactctaatt ggataccgag gggaatttat 15000
ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg accttaggcg 15060
acttttgaac gcgcaataat ggtttctgac gtatgtgctt agtcattaa actccagaaa 15120
cccgcggctg agtggctcct tcaacgttgc ggttctgtca gttccaaacg taaaacggct 15180
tgtcccgct catcggcggg ggtcataacg tgactccctt aattctccgc tcatgatcag 15240
attgtcgttt cccgccttca gtttaaacta tcagtgtttg acaggatata ttggcgggta 15300
aacctaagag aaaagagcgt ttattagaat aatcgatat taaaagggc gtgaaaagg 15360
ttatccgttc gtccatttgt atgtgcatgc caaccacagg gttccccaga tctggcgccg 15420
gccagcgaga cgagcaagat tggccgcgcg ccgaaacgat ccgacagcgc gccagcaca 15480
ggtgcgcagg caaattgcac caacgcatac agcgccagca gaatgccata gtgggcggtg 15540
acgtcgttcg agtgaaccag atcggcgagg aggcccggca gcaccggcat aatcaggccg 15600
atgccgacag cgtcgagcgc gacagtgtc agaattacga tcaggggtat gttgggtttc 15660
acgtctggcc tccggaccag cctccgctgg tccgattgaa cgcgcggtt ctttatcact 15720
gataagttgg tggacatatt atgtttatca gtgataaagt gtcaagcatg acaaagttgc 15780
agccgaatac agtgatccgt gccgccctgg acctgttgaa cgaggtcggc gtagacggtc 15840
tgacgacacg caaactggcg gaacggttgg gggttcagca gccggcgctt tactggcact 15900
tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc catgctggcg gagaatcata 15960

cgcattcgggt gccgagagcc gacgacgact ggcgctcatt tctgatcggg aatgcccgca 16020
gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg catccatgcc ggcacgcgac 16080
cgggcgaccc gcagatggaa acggccgacg cgcagcttcg ctctctctgc gaggcggggtt 16140
tttcggccgg ggacgccgtc aatgcgctga tgacaatcag ctacttcact gttggggccg 16200
tgcttgagga gcaggccggc gacagcgatg ccggcgagcg cggcggcacc gttgaacagg 16260
ctccgctctc gccgctgttg cgggcccgcga tagacgcctt cgacgaagcc ggtccggacg 16320
cagcgttcga gcagggactc gcggtgattg tcgatggatt ggcgaaaagg aggctcgttg 16380
tcaggaacgt tgaaggaccg agaaagggtg acgattgatc aggaccgctg ccggagcgca 16440
accactcac tacagcagag ccatgtagac aacatcccct ccccctttcc accgcgtcag 16500
acgcccgtag cagcccgcta cgggcttttt catgccctgc cctagcgtcc aagcctcacg 16560
gccgcgctcg gcctctctgg cggccttctg gcgctcttcc gcttcctcgc tcaactgactc 16620
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 16680
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 16740
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 16800
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgacag gactataaag 16860
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 16920
taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgcttttcc gctgcataac 16980
cctgcttcgg ggtcattata gcgatttttt cggatatatc atcctttttc gcacgatata 17040
caggattttg ccaaagggtt cgtgtagact ttctttggtg tatccaacgg cgtcagccgg 17100
gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc cttcttcact gtcccttatt 17160
cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg ctggccggct accgccggcg 17220

taacagatga gggcaagcgg atggctgatg aaaccaagcc aaccaggaag ggcagccac 17280
ctatcaaggt gtactgcctt ccagacgaac gaagagcgat tgaggaaaag gcggcgggcg 17340
ccggcatgag cctgtcggcc tacctgctgg ccgctggcca gggctacaaa atcacgggcg 17400
tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa tggcgacctg ggccgcctgg 17460
gcggcctgct gaaactctgg ctcaccgacg acccgcgcac ggcgcggttc ggtgatgcca 17520
cgatcctcgc cctgctggcg aagatcgaag agaagcagga cgagcttggc aaggatcatga 17580
tgggcgtggt ccgcccgagg gcagagccat gactttttta gccgctaaaa cggccggggg 17640
gtgcgcgtga ttgccaagca cgtcccatg cgctccatca agaagagcga cttcgcgag 17700
ctggtgaagt acatcaccga cgagcaaggc aagaccgagc gcctttgcga cgctca 17756

<210> 48

<211> 17118

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 48

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttccacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gagggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatat ttataaccttg	540
tttttaacca gggctgcgcc ctgtgcgctg gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccagg gctgcgcccc	660
tcggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgga cgcccggaag cattgacgtg	780
ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg	840
ggtggcgggc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg	900
gcaattttta ccttgggcat tcttggcata gtggtcgcg gtgccgtgct cgtgttcggg	960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa	1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag	1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata	1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaagtcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacacgcgac ttccgtccca gccgtgccag gtgctgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560

ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620

gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680

gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740

tgtatgcgcg aggttaccga ctgcggcctg agtttttttaa gtgacgtaaa atcgtgttga 1800

ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg gccatatcaa 1860

tgattttctg gtgcgtaccg ggttgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920

tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980

ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat caccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggtt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaagggtata taagctggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180
ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240
cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360
gagggatgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccgggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatgtcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gtcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgcgggcggg ttatttgcgc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttcgctatt 5100

ctggagcttg ttgtttatatt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtccctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400

cgctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc ttggttccg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcattcaggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttctc agcggcttta 5760

tccagcgatt tctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgcct tacaacggct ctcccgtga cgcggtccc gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctggt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc caggggtttt ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtagcc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatatcc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgagggtt 7440

tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaadc 7920

tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggt catgacttga gcctaaaadc cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaactctatt gcatcatcgg 8280

agaatatgga gttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgtatga aaccggaaaag gccgctcagg agctggccag cggcgcagac 8580

cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtcacctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa cttttgttgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac tttgtacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctaccccgc 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctctgtcttt 9060

cagcttcgat gtagggggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

ctacaaagat cgttatgttt atcggcactt tgcacgggc gcgctccga ttccggaagt 9180

gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240

tgtcacgttg caagacctgc ctgaaaccga actgcccgtc gttctgcagc cggtcgcgga 9300

ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg 9360

accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgagggc 9480

tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggtc tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cgccgctctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgccccca gcactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260
cttncaggct cgggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgcctctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500
tcttcttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160

cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220

ttgattacac tctaaatctc gatgaccga ccaaaaagag cgaacaaaga aataatcttg 11280

tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340

tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400

tatgatccag ttagaacaac cactcagtca tcaagcaaaa ctgactccag tactgagaag 11460

taaatctcag ttaaggggc ttttcattgc tattgtcatt gttagcgcac gggtcattag 11520

cctgagttta ttactttccc ttgacatctc aaagctaaaa ttttggatgt tattgcctgt 11580

tatactatgg caaacatttt tatatacggg attatttatt acatctcatg atgccatgca 11640

tggcgtagta tttcccaaaa acaccaagat taatcatttg attggaacat tgaccctatc 11700

cctttatggc cttttaccat atcaaaaact attgaaaaaa cattgggttac accaccacaa 11760

tccagcaagc tcaatagacc cggattttca caatggtaaa caccaaagtt tctttgcttg 11820

gtattttcat tttatgaaag gttactggag ttgggggcaa ataattgcgt tgactattat 11880

ttataacttt gctaaatata tactccatat cccaagtgat aatctaactt acttttgggt 11940

gctaccctcg cttttaagtt cattacaatt attctatttt ggtacttttt taccocatag 12000

tgaaccaata ggggggttatg ttcagcctca ttgtgcccaa acaattagcc gtccatattg 12060

gtggtcattt atcacgtgct atcatttttg ctaccacgag gaacatcacg aatatcctca 12120

tatttcttgg tggcagttac cagaaattta caaagcaaaa tagaagcttg gcgtaatcat 12180

ggcatagct gtttctgtg tgaaattgtt atccgctcac aattccacac aacatacag 12240

ccggaagcat aaagtgtaaa gcctgggggtg cctaagtgt gagctaactc acattaattg 12300

cggtgcgctc actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 12360

tcggccaacg cgcggggaga ggcgggttgc gtattggggc aaagacaaaa ggcgcacatt 12420

caaccgattg agggagggaa ggtaaatt gacggaaatt attcattaaa ggtgaattat 12480
caccgtcacc gacttgagcc atttgggaat tagagccagc aaaatcacca gtagcaccat 12540
taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12600
tagcgacaga atcaagtttg ccttttagcgt cagactgtag cgcgttttca tcggcatttt 12660
cggtcatagc ccccttatta gcgtttgcc a tcttttcata atcaaaatca ccggaaccag 12720
agccaccacc ggaaccgcct ccctcagagc cgccaccctc agaaccgcc a ccctcagagc 12780
caccaccctc agagccgcc ccagaaccac caccagagcc gccgccagca ttgacaggag 12840
gcccgatcta gtaacataga tgacaccgcg cgcgataatt taccctagtt tgccgcgtat 12900
attttgtttt ctatcgcgta ttaaattgat aattgcggga ctctaatacat aaaaacccat 12960
ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13020
aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13080
caaatgtttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13140
gaacgcagca agatatcgcg gtgcatctcg gtcttgcccg ggcagtcgcc gccgacgccg 13200
ttgatgtgga cgccgggccc gatcatattg tcgctcagga tcgtggcggt gtgcttgctg 13260
gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccgtgg 13320
gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaaa 13380
acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13440
aggttgggcg tcgcttggtc ggtcatttcg aacccagag tcccgctcag aagaactcgt 13500
caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13560
ggaagcggtc agccattcg ccgccaagct cttcagcaat atcacgggta gccaacgcta 13620
tgtcctgata gcggtcgcc acaccagcc ggccacagtc gatgaatcca gaaaagcggc 13680

cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13740

cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13800

cttcgtccag atcatcctga tcgacaagac cggcttccat ccgagtacgt gctcgcctga 13860

tgcgatgttt cgcttggtgg tcgaatgggc aggtagccgg atcaagcgta tgcagccgcc 13920

gcattgcacg agccatgatg gatactttct cggcaggagc aaggtgagat gacaggagat 13980

cctgccccgg cacttcgccc aatagcagcc agtcccttcc cgcttcagtg acaacgtcga 14040

gcacagctgc gcaaggaacg cccgtcgtgg ccagccacga tagccgcgct gcctcgtcct 14100

gcagttcatt cagggcacccg gacaggtcgg tcttgacaaa aagaaccggg cgccctgcg 14160

ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttggtgcc cagtcatagc 14220

cgaatagcct ctccacccaa gcggccggag aacctgcgtg caatccatct tgttcaatca 14280

tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14340

ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatttgc 14400

tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14460

ttagctcatt aaactccaga aaccgcggc tgagtggctc cttcaacgtt gcggttctgt 14520

cagttccaaa cgtaaaacgg cttgtccgc gtcacggcg ggggtcataa cgtgactccc 14580

ttaattctcc gtcatgatc agattgtcgt tccccgcctt cagtttaaac tatcagtgtt 14640

tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataatcggat 14700

atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaaaccaca 14760

gggttcccca gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg 14820

atccgacagc gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 14880

cagaatgcc a tagtgggagg tgacgtcgtt cgagtgaacc agatcgcgca ggaggccccg 14940

cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15000
gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15060
aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15120
gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15180
aacgaggtcg gcgtagacgg tctgacgaca cgaaaactgg cggaacgggt gggggttcag 15240
cagccggcgc ttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15300
gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15360
tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15420
cgcattcatg ccggcacgcg accgggcgca ccgcagatgg aaacggccga cgcgcagctt 15480
cgcttctctt gcgaggcggg tttttcggcc ggggacgccg tcaatgcgct gatgacaatc 15540
agctacttca ctgttggggc cgtgcttgag gagcaggccg gcgacagcga tgccggcgag 15600
cgcggcgcca ccgttgaaca ggctccgctc tcgccgctgt tcggggccgc gatagacgcc 15660
ttcgacgaag ccggtccgga cgcagcgctc gagcaggggac tcgcggtgat tgtcgatgga 15720
ttggcgaaaa ggaggctcgt tgtcaggaac gttgaaggac cgagaaaggg tgacgattga 15780
tcaggaccgc tgccggagcg caaccactc actacagcag agccatgtag acaacatccc 15840
ctcccccttt ccaccgctc agacgcccg agcagccgc tacgggcttt ttcatgccct 15900
gccctagcgt ccaagcctca cggccgcgct.cggcctctct ggccgccttc tggcgctctt 15960
ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag 16020
ctcactcaaa ggcggttaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16080
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgcttt 16140
tccataggct ccgccccct gacgagcatc aaaaaatcg acgctcaagt cagaggtggc 16200

gaaacccgac aggactataa agataccagg cgtttccccc tgggaagctcc ctctgtgcgt 16260

ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 16320

tggcgctttt ccgctgcata acctgcttc ggggtcatta tagcgatttt ttcggtatat 16380

ccatcctttt tcgcacgata tacaggattt tgccaaaggg ttcgtgtaga ctttccttgg 16440

tgtatccaac ggcgtcagcc gggcaggata ggtgaagtag gccacccgc gagcgggtgt 16500

tccttcttca ctgtccctta ttgcacctg gcggtgctca acgggaatcc tgctctgcga 16560

ggctggccgg ctaccgccgg cgtaacagat gagggcaagc ggatggctga tgaaaccaag 16620

ccaaccagga agggcagccc acctatcaag gtgtactgcc ttccagacga acgaagagcg 16680

attgaggaaa aggcggcggc ggccggcatg agcctgtcgg cctacctgct ggccgtcggc 16740

cagggctaca aaatcacggg cgtcgtggac tatgagcacg tccgcgagct ggcccgcac 16800

aatggcgacc tgggccgcct gggcggcctg ctgaaactct ggctcaccga cgacccgcgc 16860

acggcgcggt tcggtgatgc cacgatcctc gccctgctgg cgaagatcga agagaagcag 16920

gacgagcttg gcaaggatcat gatgggcgtg gtccgcccga gggcagagcc atgacttttt 16980

tagccgctaa aacggccggg ggggtgcgct gattgccaag cagtcacca tgcgctccat 17040

caagaagagc gacttcgcgg agctggtgaa gtacatcacc gacgagcaag gcaagaccga 17100

gcgcctttgc gacgctca 17118

<210> 49

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 49

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt	60
cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct	240
tccggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atatctctcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720

tccgatagct ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780

tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

caccacaaaa gtcagacggc gtaacaaaa gtcacacaac acaagctgta aggatttcgg 900

cacggctacg gaagacggag aagccacctt cagtggactc gagtaccatt taattctatt 960

tgtgtttgat cgagacctaa tacagccctt acaacgacca tcaaagtcgt atagctacca 1020

gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa 1080

ctatacagaa taagataggt ggagagctta taccgagctc ccaaactctgt ccagatcatg 1140

gttgaccggt gcctggatct tcctatagaa tcctccttat tcgttgacct agctgattct 1200

ggagtgacct agagggatcat gacttgagcc taaaatccgc cgcctccacc atttgtagaa 1260

aatgtgacg aactcgtgag ctctgtacag tgaccgggtga ctctttctgg catgcggaga 1320

gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380

ctctgagggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440

aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcacgggaga atatggagct 1500

tcacgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560

cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgatc tggtaaaaga 1620

ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680

aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740

tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800

tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860

ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgtc 1920

agccaacat ttgttgccat attttctcgc tctccccacc agctgctctt ttcttttctc 1980

tttcttttcc catcttcagt atattcatct tcccatcaa gaacctttat ttcccctaag 2040
taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct accccgcttg agcagacatc 2160
accatgcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctccggagggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcgctg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcgcccgcg ctcccgatcc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcac tcccgccgtg cacaggggtgt cacgttgcaa 2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggagggc catggatgcg 2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
ggccaataca ctacatggcg tgatttcata tgccgattg ctgatcccca tgtgtatcac 2640
tggcaaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga ttccggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcgggtca ttgactggag cgaggcgatg 2820
ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgagg atcgccgcgg 2940
ctccgggcgt atatgctccg cattggctctt gaccaactct atcagagctt gggtgacggc 3000
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgctccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgcccg cgaagcgcg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcggtccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240

tgacgaatct ggatataaga tcgttgggtgt cgatgtcagc tccggagttg agacaaatgg 3300

tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagtg 3360

atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420

cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480

gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540

tcaacgggtcg tcaagagacc tacgagactg aggaatccgc tcttgggtcc acgcgactat 3600

atatttgtct ctaattgtac ttgacatgc tcctcttctt tactctgata gcttgactat 3660

gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780

gatggtatac aatagtaacc atgcatgggt gcctagttaa tgctccgtaa caccacaatac 3840

gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggtaca 3900

cttgttttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960

ctccacatct ccactcgacc tgcaggcatg caaagcttga gattaaaata gataaggaaa 4020

agaaagttaa aagaaattcg gaagcatggc acattcttct ttttataaat acatgcctga 4080

ctttcttttt ccatcgatat gatatatgca tatgatagat atacaagcaa tcttcttcaa 4140

ggagtttgaa attttgtcct ccaggagcaa aaaaaagttt tttttatac atgtttgtac 4200

acaagaatag ttaccaattt gctttgggtct tacgtgctgc aagtttatat cgttttcaat 4260

ttctttgtct ttacattttc ttgttccttt atctttcctc atttagtctt tgggagaatt 4320

aggaaaaggg agcggaaagg taagaaatgc ttgcgtattt tactaattcg gcaaacatcc 4380

aatttggcaa acagcagcct gtgcaacgct ctcgagatga cagtatcttt gattacactc 4440

taa atctcga tgacccgacc aaaaagagcg aacaaagaaa taatcttggtg cattcgaata 4500

tgatggaaga ttttttcccc cttattctaa atgttgacat agcgtgtatg ttatataaac 4560

aaaaagaaat tgtacaaact ttcttttctt ctctttttat tttatctcta tgctgtcgaa 4620

gctgcagtca atcagcgtca aggcccgccg cgttgaacta gcccgggaca tcacgcggcc 4680

caaagtctgc ctgcatgctc agcgggtgctc gttagttcgg ctgcgagtgg cagcaccaca 4740

gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800

tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaacggga 4860

gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920

cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgag tgccatgggg 4980

tgaagtggct ggcactctcc tcttggtggt tgggtggcgcg ctccggcatgg agatgtatgc 5040

ccgctatgca cacaaagcca tctggcatga gtcgcctctg ggctggctgc tgcacaagag 5100

ccaccacaca cctcgactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5160

actgcccgcc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tggggggcggc 5220

ctgctttgga gcgggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280

tggcctgggtg cacaggcgct ttcccaccgg gcccatcgct ggctggcctt acatgaagcg 5340

cctgacagtg gccaccagc tacaccacag cggcaagtac ggtggcgcg cctgggggtat 5400

gttcttgggt ccacaggagc tgcagcacat tccagggtcg gcggaggagg tggagcgact 5460

ggtcctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520

atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5580

cgcccgtagc ctaggtgatc aaagactgtt gggcttgtct ctgaagcttg taggaaaggc 5640

agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700

cggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa 5760
gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaaga ttgctctgtt 5820
gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880
gattaagagt ttttttggtg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940
aaagagacaa cacatacaat ataaaatcat tgtctttaaa tgtaattta gtagagtga 6000
aagcctgcat ttttttgta cgcataaaca atgaattcac cccgcttctg gtttttaaat 6060
aattatgtca aactagggaa aattcttttt tttctcttcg ttcttttttt ggcttggtgt 6120
ggagtcacag gcttgcttc agattgatag aggttgata cactcaacag agcaatcttg 6180
gcacgttcgc ttcttttag atgagctctt gtaggatatg ttgtctctg ctgcttgagt 6240
acagatccaa ttgcagcata tacttgga gacgctcgta caccgccttg acagtttgcc 6300
ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc 6360
ttcagagaca agcccaacag tctttgatca ctaggctac gggcggttgcc ttgctgtatt 6420
tgttctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcctc 6480
gagccatact acagcataaa aggatacgtt ttctttaaca gaaatttacc cttttgttat 6540
cagcacatac aaaaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaattt 6600
tattcaatcc ataaatgaat ttttttgga caaaaaagaa agattatgcc tgattttctc 6660
tatttttttt ttttttacia ctccaccaat actttctagc ccagcttggc gtaatcatgg 6720
tcatagctgt ttctgtgtg aaattgttat ccgctcacia ttccacacia catagagcc 6780
ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac attaattgcg 6840
ttgcgctcac tgcccgttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 6900
ggccaacgcg cggggagagg cggtttgcgt attgggcaa agacaaaagg gcgacattca 6960

accgattgag ggaggggaagg taaatattga cggaaattat tcattaaagg tgaattatca 7020

ccgtcaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta 7080

ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta 7140

gcgacagaat caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg 7200

gtcatagccc ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag 7260

ccaccaccgg aaccgcctcc ctacagagccg ccaccctcag aaccgccacc ctacagagcca 7320

ccaccctcag agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc 7380

ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat 7440

tttgttttct atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct 7500

cataaataac gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa 7560

ttatatgata atcatcgcaa gaccggcaac aggattcaat ctttaagaaac tttattgcca 7620

aatgtttgaa cgatcgggga tcatccgggt ctgtggcggg aactccacga aaatatccga 7680

acgcagcaag atatcgcggt gcatctcggt cttgcctggg cagtcgccgc cgacgccgtt 7740

gatgtggacg ccgggcccga tcatattgtc gctcaggatc gtggcgttgt gcttgctggc 7800

cgttgctgtc gtaatgatat cggcaccttc gaccgcctgt tccgagaga tcccgtgggc 7860

gaagaactcc agcatgagat ccccgcgctg gaggatcatc cagccggcgt cccggaaaac 7920

gattccgaag cccaaccttt catagaaggc ggcggtggaa tcgaaatctc gtgatggcag 7980

gttgggcgtc gcttggtcgg tcatttcgaa cccagagtc ccgctcagaa gaactcgtca 8040

agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100

aagcggtcag ccatttcgcc gccaaagtct tcagcaatat cacgggtagc caacgctatg 8160

tcctgatagc ggtccgccac acccagccgg ccacagtcga tgaatccaga aaagcggcca 8220

ttttccacca tgatattcgg caagcaggca tcgcatggg tcacgacgag atcatcgccg 8280

tcgggcatgc gcgccttgag cctggcgaac agttcggctg gcgcgagccc ctgatgctct 8340

tcgtccagat catcctgac gacaagaccg gcttccatcc gagtacgtgc tcgtcogatg 8400

cgatgtttcg cttgggtggc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8460

attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8520

tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc 8580

acagctgcgc aaggaacgcc cgtcgtggcc agccacgata gccgcgctgc ctcgtcctgc 8640

agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8700

gacagccgga acacggcggc atcagagcag ccgattgtct gttgtgcca gtcatagccg 8760

aatagcctct ccaccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8820

cgaaacgac cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8880

ggataccgag ggggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 8940

gctgatagt accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9000

agctcattaa actccagaaa cccgcggctg agtggctcct tcaacgttgc ggttctgtca 9060

gttccaaacg taaaacggct tgtcccgct catcgccggg ggtcataacg tgactccctt 9120

aattctccgc tcatgatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 9180

acaggatata ttggcgggta aacctagag aaaagagcgt ttattagaat aatcgatat 9240

ttaaaagggc gtgaaaagg ttatccgttc gtccatttgt atgtgcatgc caaccacagg 9300

gttccccaga tctggcgccg gccagcgaga cgagcaagat tggccgccgc ccgaaacgat 9360

ccgacagcgc gccagcaca ggtgcgcagg caaattgcac caacgcatac agcgccagca 9420

gaatgccata gtgggcgggtg acgtcgttcg agtgaaccag atcgcgcagg agggccggca 9480

gcaccggcat aatcaggccg atgccgacag cgtcgagcgc gacagtgctc agaattacga 9540
tcaggggtat gttgggtttc acgtctggcc tccggaccag cctccgctgg tccgattgaa 9600
cgcgcggtatt ctttatcact gataagttgg tggacatatt atgtttatca gtgataaagt 9660
gtcaagcatg acaaagttgc agccgaatac agtgatccgt gccgccctgg acctgttgaa 9720
cgaggtcggc gtagacggtc tgacgacacg caaactggcg gaacggttgg gggttcagca 9780
gccggcgctt tactggcact tcaggaacaa gcgggcgctg ctcgacgcac tggccgaagc 9840
catgctggcg gagaatcata cgcattcggg gccgagagcc gacgacgact ggcgctcatt 9900
tctgatcggg aatgcccga gcttcaggca ggcgctgctc gcctaccgcg atggcgcgcg 9960
catccatgcc ggcacgcgac cgggcgcacc gcagatggaa acggccgacg cgcagcttcg 10020
cttcctctgc gaggcgggtt ttctggccgg ggacgccgctc aatgcgctga tgacaatcag 10080
ctacttcact gttggggccg tgcttgagga gcaggccggc gacagcgatg ccggcgagcg 10140
cggcggcacc gttgaacagg ctccgctctc gccgctgttg cgggccgcga tagacgcctt 10200
cgacgaagcc ggtccggacg cagcgttcga gcagggactc gcggtgattg tcgatggatt 10260
ggcgaaaagg aggctcgttg tcaggaacgt tgaaggaccg agaaaggggtg acgattgatc 10320
aggaccgctg ccggagcgca acccactcac tacagcagag ccatgtagac aacatcccct 10380
ccccctttcc accgcgtcag acgcccgtag cagcccgcta cgggcttttt catgccctgc 10440
cctagcgtcc aagcctcacg gccgcgctcg gcctctctgg cggccttctg gcgctcttcc 10500
gcttctctgc tactgactc gctgcgctcg gtcgttcggc tgccggcgagc ggtatcagct 10560
cactcaaagg cggtaatcag gttatccaca gaatcagggg ataacgcagg aaagaacatg 10620
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttget ggcgtttttc 10680
cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 10740

aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 10800
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc ggggaagcgtg 10860
gcgcttttcc gctgcataac cctgcttcgg ggtcattata gcgatttttt cggatatatcc 10920
atcctttttc gcacgatata caggattttg ccaaagggtt cgtgtagact ttccttggtg 10980
tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccaccgcga gcgggtgttc 11040
cttcttcact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 11100
ctggccggct accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 11160
aaccaggaag ggcagccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11220
tgaggaaaag gcggcgccgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 11280
gggctacaaa atcacggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 11340
tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 11400
ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 11460
cgagcttggc aaggtcatga tgggcgtggt ccgcccgagg gcagagccat gactttttta 11520
gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtcccatg cgctccatca 11580
agaagagcga cttcgcgag ctggtgaagt acatcaccga cgagcaaggc aagaccgagc 11640
gcctttgcga cgctaccgg gctggttgcc ctgcgcgtg ggctggcggc cgtctatggc 11700
cctgcaaacg cgccagaaac gccgtogaag ccgtgtgcga gacaccgcgg ccgcccgcgt 11760
tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 11820
tgaggggccg actcaccgg cgcggcgttg acagatgagg ggcaggctcg atttcggccg 11880
gcgacgtgga gctggccagc ctgcgaaatc ggcgaaaacg cctgatttta cgcgagtttc 11940

ccacagatga tgtggacaag cctggggata agtgcctgc ggtattgaca cttgaggggc 12000
gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 12060
tgaggggagc acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 12120
aagggtttcc gcccgttttt cgccaccgc taacctgtct tttaacctgc ttttaaacca 12180
atatttataa accttgtttt taaccagggc tgcgcctgt gcgcgtgacc gcgcacgccg 12240
aaggggggtg ccccccttc tcgaaccctc ccggcccgct aacgcgggcc tcccatcccc 12300
ccaggggctg cgccttcg cgcgaacgg cctcaccca aaaatggcag cgctggcagt 12360
ccttgccatt gccgggatcg gggcagtaac gggatgggcg atcagcccga gcgcgacgcc 12420
cggaagcatt gacgtgccgc aggtgctggc atcgacattc agcgaccagg tgccgggcag 12480
tgagggcggc ggcctgggtg gcggcctgcc cttcacttcg gccgtcgggg cattcacgga 12540
cttcatggcg gggccggcaa tttttacctt gggcattctt ggcatagtgg tcgcgggtgc 12600
cgtgctcgtg ttcgggggtg cgataaacc agcgaaccat ttgaggtgat aggtaagatt 12660
ataccgaggt atgaaaacga gaattggacc ttacagaat tactctatga agcgccatat 12720
ttaaaaagct accaagacga agaggatgaa gaggatgagg aggcagattg ctttgaatat 12780
attgacaata ctgataagat aatatatctt ttatatagaa gatatcgccg tatgtaagga 12840
tttcaggggg caaggcatag gcagcgcgt tatcaatata tctatagaat gggcaaagca 12900
taaaaacttg catggactaa tgcttgaaac ccaggacaat aaccttatag cttgtaaatt 12960
ctatcataat tgggtaatga ctccaactta ttgatagtgt tttatgttca gataatgccc 13020
gatgactttg tcatgcagct ccaccgattt tgagaacgac agcgacttcc gtcccagccg 13080
tgccaggtgc tgctcagat tcaggttatg ccgctcaatt cgctgcgtat atcgcttgct 13140
gattacgtgc agctttccct tcaggcggga ttcatacagc ggccagccat ccgtcatcca 13200

tatcaccacg tcaaaggggtg acagcaggct cataagacgc cccagcgtcg ccatagtgcg 13260

ttcaccgaat acgtgcgcaa caaccgtctt cgggagactg tcatacgcgt aaaacagcca 13320

gcgctggcgc gatttagccc cgacatagcc ccactgttcg tccatttccg cgcagacgat 13380

gacgtcactg cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga 13440

cgtaaaatcg tgttgaggcc aacgcccata atgcgggctg ttgcccggca tccaacgcca 13500

ttcatggcca tatcaatgat tttctgggtgc gtaccggggtt gagaagcggg gtaagtgaac 13560

tgcagttgcc atgttttacg gcagtgcgag cagagatagc gctgatgtcc ggcgggtgctt 13620

ttgccgttac gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa 13680

gccactggag cacctcaaaa acaccatcat aactaaatc agtaagttgg cagcatcacc 13740

cataattgtg gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac 13800

aactttgaaa aagctgtttt ctgggtattta aggttttaga atgcaaggaa cagtgaattg 13860

gagttcgtct tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa 13920

ggaaataata aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat 13980

accgctgcgt aaaagatacg gaaggaatgt ctctgctaa ggtatataag ctggtgggag 14040

aaaatgaaaa cctatatatta aaaatgacgg acagccggta taaagggacc acctatgatg 14100

tggaacggga aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaagggtcc 14160

tgcactttga acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc 14220

tttgctcgga agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg 14280

agtgcacag gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag 14340

acagccgctt agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg 14400

aaaactggga agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga 14460

cggaaaagcc cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct 14520
ttgtgaaaga tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca 14580
agtggatatga cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt 14640
atgtcgagct attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt 14700
atattttact ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag 14760
caggagcgcga ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc 14820
aagtatttgg gcaaggggtc gctgggtattc gtgcagggca agattcggaa taccaagtac 14880
gagaaggacg gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg 14940
gacaccaagg caccaggcgg gtcaaatacag gaataagggc acattgcccc ggcgtgagtc 15000
ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa 15060
gaactgatcg acgcgggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc 15120
atgcgtgcgc cccgcgaaac cttccagtcc gtcggctcga tgggccagca agctacggcc 15180
aagatcgagc gcgacacggt gcaactggct cccctgccc tgcccgcgcc atcggccgcc 15240
gtggagcggt cgcgtcgtct cgaacaggag gcggcagggt tggcgaagtc gatgaccatc 15300
gacacgcgag gaactatgac gaccaagaag cgaaaaaccg ccggcgagga cctggcaaaa 15360
caggtcagcg aggccaagca ggccgcgttg ctgaaacaca cgaagcagca gatcaaggaa 15420
atgcagcttt ccttgttcga tattgcgccg tggccggaca cgatgcgagc gatgccaaac 15480
gacacggccc gctctgccct gttcaccacg cgcaacaaga aaatcccgcg cgaggcgctg 15540
caaaacaagg tcattttcca cgtcaacaag gacgtgaaga tcacctacac cggcgctcgag 15600
ctgcgggccc acgatgacga actggtgtgg cagcagggtg tggagtacgc gaagcgcacc 15660
cctatcggcg agccgatcac cttcacgttc tacgagcttt gccaggacct ggcgtggctg 15720

atcaatggcc ggtattacac gaaggccgag gaatgcctgt cgcgcctaca ggcgacggcg 15780
atgggcttca cgtccgaccg cgttgggcac ctggaatcgg tgtegtgtgt gcaccgcttc 15840
cgcgtcctgg accgtggcaa gaaaacgtcc cgttgccagg tcctgatcga cgaggaaatc 15900
gtcgtgtgtgt ttgctggcga ccactacacg aaattcatat gggagaagta ccgcaagctg 15960
tcgccgacgg cccgacggat gttcgactat ttcagctcgc accgggagcc gtaccgctc 16020
aagctggaaa ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 16080
gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctgggtgga acacgcctgg 16140
gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200
ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260
tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag 16320
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500
tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggtcttc aacaggagg 16560
acggcccaa ggacgtcac aaggcgcac tcgtccggcgt ttcgtggag cccgaacagc 16620
gaggccgagg ggtcgccggt atgctgctgc gggcggtgcc ggcgggttta ttgctcgtga 16680
tgatcgctcc acagattcca acgggaatct ggtggatgcg catcttcac ctcgggcac 16740
ttaatatttc gctattctgg agcttggtgt ttatttcggt ctaccgcctg ccgggcgggg 16800
tcggggcgac ggtaggcgct gtgcagccgc tgatggctgt gttcatctct gccgctctgc 16860
taggtagccc gatacgattg atggcggctc tgggggctat ttgcggaact gcgggcgtgg 16920
cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgctcga gcgggcctgg 16980

cgggggcggt ttccatggcg ttcggaaccg tgctgaccg caagtggcaa cctcccgtgc 17040
ctctgtcac ctttaccgcc tggcaactgg cggccggagg acttctgtc gttccagtag 17100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt 17160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220
tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 17280
cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag 17340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17400
ttcctcagcg gctttatcca gcgatttctt attatgtcgg catagtctc aagatcgaca 17460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17580
ccgcgagatc atccgtgttt caaaccggc agcttagttg ccgttcttcc gaatagcatc 17640
ggtaacatga gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact 17700
gatgggctgc ctgtatcgag tggtgathtt gtgccgagct gccggtcggg gagctgttgg 17760
ctggctggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17820
acattgcgga cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa 17880
cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggg ccacgctggg 17940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18000
aaatcaaaag aatagcccga gatagggttg agtgtgttc cagtttgga caagagtcca 18060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18120
ccactacgtg aaccatcacc caaatcaagt tttttggggg cgaggtgccg taaagcacta 18180
aatcggaacc ctaaaggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg 18240

gcgagaaagg aaggggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg 18300
ggaagggcgga tcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc 18360
tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 18420
ggccagtga ttcgagctcg gtacccggg 18449

<210> 50

<211> 18617

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 50

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtgga tacctcgcg 120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180
ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240

cagcctcgca aatcgggcgaa aacgcctgat tttagcgag tttccacag atgatgtgga 300

caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360

gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggcca cgcctaact gtcttttaac ctgcttttaa accaatattt ataaaccttg 540

tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600

cttctcgaac cctcccggcc cgctaacgcg ggctcccat cccccaggg gctgcgcccc 660

tcggccgga acggcctcac ccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720

atcggggcag taacgggatg ggcgatcagc ccgagcgga cggccggaag cattgacgtg 780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840

ggtggcgggc tgcccttcac ttccggcgtc ggggcattca cggacttcat ggcggggccg 900

gcaattttta ccttgggcat tcttggcata gtggtcgagg gtgccgtgct cgtgttcggg 960

ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080

acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140

agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200

ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

ctaattgctt aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320

atgactccaa cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc 1380

agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgtgcctc 1440

agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500

cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560
ggtgacagca ggctcataag acgccccagc gtgcgcatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatat gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg gggtgagaag cgggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgtgat gtccggcggg gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100
aaatcggctc cgtcgatact atgttatatc ccaactttga aaacaacttt gaaaaagctg 2160
ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagtgc gtcttgttat 2220
aattagcttc ttggggatat tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340
tacggaagga atgtctcctg ctaaggtata taagctgggt ggagaaaatg aaaacctata 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacggca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc tcaaggctctt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggg atgacattgc 2940

cttctcgctc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggg attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360

gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420

ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480

aaaccttcca gtccgtcggc tcgatgggtc agcaagctac ggccaagatc gagcgcgaca 3540

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600

gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660

tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720

agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780

tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840

ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900

tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960

acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc ggcgagccga 4020

tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080

acacgaaggc cgaggaatgc ctgtcgcgc tacaggcgac ggcgatgggc ttcacgtccg 4140

accgcgttgg gcacctggaa tcgggtgtgc tgctgcaccg cttccgcgtc ctggaccgtg 4200

gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg ctgtttgctg 4260

gcgaccacta cacgaaattc atatggggaga agtaccgcaa gctgtcgccg acggcccgcac 4320

ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380

gcctcatgtg cggatcggat tccaccgcgc tgaagaagtg gcgcgagcag gtcggcgaag 4440

cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500

tgcatctcaa acgctagggc cttgtggggc cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620

gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680

tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740

ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800

gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4860

catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920

tcacaaggcg catctgtccg gcgttttcgt ggagcccga cagcgaggcc gaggggtcgc 4980

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040

tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgtatt 5100

ctggagcttg ttgtttattt cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220

attgatggcg gtcctggggg ctatthtcgg aactgcgggc gtggcgctgt tgggtgtgac 5280

accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340

ggcgttcgga accgtgctga cccgcaagt gcaacctccc gtgcctctgc tcacctttac 5400

cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460

gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520

agcgggttta acctacttcc tttggttcgg ggggatctcg cgactcgaac ctacagttgt 5580

ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacggc 5640

cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700

tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760

tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cagggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgcgcct tacaacggct ctcccgtga cgcgctccg gactgatggg ctgcctgtat 6060

cgagtgggta ttttgtccg agctgccggg cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatgggtgt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcgggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcgggtaccg ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactaccaa catggagtag gataagggcc agttccgcca 6900

gtcattaag agccagttca tgggcgttgg catgatggcc gtcattgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgctccct taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttgggt ttcaatatc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440

tcgaaataca tccgatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaate 7920

tgtccagatc atgggtgacc ggtgcctgga tcttcctata gaatcatcct tattegttga 7980

cctagctgat tctggagtga ccagaggggt catgacttga gcctaaaate cgccgcctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgagg agagacggac ggacgcagag agaagggtg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340

gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccgggc 8460

gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520

ctgctttgcc cgggtgatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580

cggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgagggtccc 8640

tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700

aggtcgttgc gtcagtccaa catttggtgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820

tatttcccct aagtaagtac ttgtctacat ccatactcca tccttcccat cccttattcc 8880

tttgaacctt tcagttcgag ctttcccact tcatcgagc ttgactaaca gctaccccg 8940

ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000

aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgtt gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgtg cggccgatct tagccagacg agcgggttcg gccattcgg 9360
accgcaagga atcgggtcaat aactacatg gcgtgatttc atatgcgca ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660
gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720
aggatcgccg cggtccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900
gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080
ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200
cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380
tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440
atagcttgac tatgaaaatt ccgtcaccag cncctggggt cgcaaagata attgcatggt 10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560
canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttaccaga 10680
atgcacaggt acacttggtt agaggtaatc cttcttttcta gctagaagtc ctcgtgtact 10740
gtgtaagcgc cactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800
tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgaccga ccaaaaagag cgaacaaaga aataatcttg 11280
tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340
tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400
tatgctgtcg aagctgcagt caatcagcgt caaggcccgc cgcgttgaac tagcccgcga 11460

catcacgagg cccaaagtct gcctgcatgc tcagcgggtgc tcgttagttc ggctgagagt 11520

ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgccggcg cgggcgatga 11580

gcacagcgcc gatgtagcac tccagcagct tgaccgggct atcgagagc gtcgtgcccg 11640

gcgcaaacgg gagcagctgt cataccaggc tgccgccatt gcagcatcaa ttggcgtgtc 11700

aggcattgcc atcttcgcca cctacctgag atttgccatg cacatgaccg tggggcggcg 11760

agtgccatgg ggtgaagtgg ctggcactct cctcttggtg gttgggtggcg cgctcggcat 11820

ggagatgtat gcccgtatg cacacaaagc catctggcat gagtcgcctc tgggctggct 11880

gctgcacaag agccaccaca cacctcgac tggacccttt gaagccaacg acttgtttgc 11940

aatcatcaat ggactgcccg ccatgctcct gtgtaccttt ggcttctggc tgcccaacgt 12000

cctggggggcg gcctgctttg gagcggggct gggcatcacg ctatacggca tggcatatat 12060

gtttgtacac gatggcctgg tgcacaggcg ctttcccacc gggcccatcg ctggcctgcc 12120

ctacatgaag cgctgacag tggcccacca gctacaccac agcggcaagt acggtggcg 12180

gccctgggggt atgttcttgg gtccacagga gctgcagcac attccaggtg cggcggagga 12240

ggtggagcga ctggctctgg aactggactg gtccaagcgg tagaagcttg agattaaaat 12300

agataaggaa aagaaagtga aaagaaattc ggaagcatgg cacattcttc tttttataaa 12360

tacatgcctg actttctttt tccatcgata tgatatatgc atatgataga tatacaagca 12420

atcttcttca aggagtgtga aattttgtcc tccaggagca aaaaaagtt tttttttata 12480

catgtttgta cacaagaata gttaccaatt tgctttggtc ttacgtgctg caagtttata 12540

tcgttttcaa tttctttgtc ttacattttt ctttgcctt tatctttcct catttagtct 12600

ttgggagaat taggaaaagg gagcggaaag gtaagaaatg cttgcgtatt ttactaatc 12660

ggcaaacatc caatttggca aacagcagcc tgtgcaacgc tctcgagatg acagtatctt 12720

tgattacact ctaaattctcg atgacccgac caaaaagagc gaacaaagaa ataattctgt 12780

gcattcgaat atgatggaag attttttccc ccttattcta aatggtgaca tagcgtgtat 12840

gttatataaa caaaaagaaa ttgtacaaac tttcttttct tctcttttta ttttatctct 12900

atgatccagt tagaacaacc actcagtcac caagcaaac tgactccagt actgagaagt 12960

aaatctcagt ttaaggggct tttcattgct attgtcattg ttagcgcagtg ggtcattagc 13020

ctgagtttat tactttccct tgacatctca aagctaaaat tttggatgtt attgcctgtt 13080

atactatggc aaacattttt atatacggga ttatttatta catctcatga tgccatgcat 13140

ggcgtagtat ttccccaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 13200

ctttatggtc ttttaccata tcaaaaacta ttgaaaaaac attgggttaca ccaccacaat 13260

ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg 13320

tattttcatt ttatgaaagg ttactggagt tgggggcaaa taattgcgtt gactattatt 13380

tataactttg ctaaatacat actccatctc ccaagtgata atctaactta cttttgggtg 13440

ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt 13500

gaaccaatag ggggttatgt tcagcctcat tgtgcccaaa caattagccg tcctatttgg 13560

tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atatcctcat 13620

atttcttggg ggcagttacc agaaatttac aaagcaaaat agaagcttgg cgtaatcatg 13680

gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc 13740

cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg agctaactca cattaattgc 13800

gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc attaataaat 13860

cggccaacgc gcggggagag gcggtttgcg tattgggcca aagacaaaag ggcgacattc 13920

aaccgattga gggagggaag gttaaatttg acggaaatta ttcattaaag gtgaattatc 13980

accgtcaccg acttgagcca tttgggaatt agagccagca aaatcaccag tagcaccatt 14040

accattagca aggccgaaa cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt 14100

agcgacagaa tcaagtttgc ctttagcgtc agactgtagc gcgttttcat cggcattttc 14160

ggtcatagcc cccttattag cgtttgccat cttttcataa tcaaaatcac cggaaccaga 14220

gccaccaccg gaaccgcctc cctcagagcc gccaccctca gaaccgccac cctcagagcc 14280

accaccctca gagccgccac cagaaccacc accagagccg ccgccagcat tgacaggagg 14340

cccgatctag taacatagat gacaccgcgc gcgataattt atcctagttt gcgcgctata 14400

ttttgttttc tatcgcgat taaatgtata attgcgggac tctaatacata aaaacccatc 14460

tcataaataa cgtcatgcat tacatgttaa ttattacatg cttaacgtaa ttcaacagaa 14520

attatatgat aatcatcgca agaccggcaa caggattcaa tcttaagaaa ctttattgcc 14580

aaatgtttga acgatcgggg atcatccggg tctgtggcgg gaactccacg aaaatatccg 14640

aacgcagcaa gatatcgcgg tgcattctcg tcttgccctgg gcagtcgccg ccgacgccgt 14700

tgatgtggac gccggggccc atcatattgt cgctcaggat cgtggcggtg tgcttgctcg 14760

ccgttgctgt cgtaatgata tcggcacctt cgaccgcctg ttccgcagag atcccgtggg 14820

cgaagaactc cagcatgaga tccccgcgct ggaggatcat ccagccggcg tccccgaaaa 14880

cgattccgaa gcccaacctt tcatagaagg cggcggtgga atcgaaatct cgtgatggca 14940

ggttgggcgt cgcttggtcg gtcatttcga accccagagt cccgctcaga agaactcgtc 15000

aagaaggcga tagaaggcga tgcgctgcga atcgggagcg gcgataccgt aaagcacgag 15060

gaagcgggtca gcccatctgc cgccaagctc ttcagcaata tcacgggtag ccaacgctat 15120

gtcctgatag cggtcgcgca caccagccg gccacagtcg atgaatccag aaaagcggcc 15180

attttccacc atgatattcg gcaagcaggc atcgccatgg gtcacgacga gatcatcgcc 15240

gtcgggcatg cgcgccttga gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc 15300
ttcgtccaga tcctcctgat cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat 15360
gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgtat gcagccgccg 15420
cattgcatca gccatgatgg atactttctc ggcaggagca aggtgagatg acaggagatc 15480
ctgccccggc acttcgcca atagcagcca gtcccttccc gcttcagtga caacgtcgag 15540
cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg 15600
cagttcattc agggcaccgg acaggtcggt cttgacaaaa agaaccgggc gccctgcgc 15660
tgacagccgg aacacggcgg catcagagca gccgattgtc tgttggtgcc agtcatagcc 15720
gaatagcctc tccaccaag cggccggaga acctgcgtgc aatccatctt gttcaatcat 15780
gcgaaacgat ccagatccgg tgcagattat ttggattgag agtgaatatg agactctaata 15840
tggataccga ggggaattta tggaacgtca gtggagcatt ttgacaaga aatatttgct 15900
agctgatagt gacctaggc gacttttgaa cgcgcaataa tggtttctga cgtatgtgct 15960
tagctcatta aactccagaa acccgcggtc gagggtcc ttcaacgttg cggttctgtc 16020
agttccaaac gtaaaacggc ttgtcccgcg tcctcggcgg gggtcataac gtgactccct 16080
taattctccg ctcatgatca gattgtcggt tccgccttc agtttaaact atcagtgttt 16140
gacaggatat attggcgggt aaacctaaga gaaaagagcg tttattagaa taatcggata 16200
tttaaaaggc cgtgaaaagg tttatccgtt cgtccatttg tatgtgcatg ccaaccacag 16260
ggttccccag atctggcgcc ggccagcgag acgagcaaga ttggccgccg cccgaaacga 16320
tccgacagcg cgccagcac aggtgcgcag gcaaattgca ccaacgcata cagcgccagc 16380
agaatgccat agtgggcggg gacgtcgttc gagtgaacca gatcgcgcag gaggcccgcc 16440
agcaccggca taatcaggcc gatgccgaca gcgtcgagcg cgacagtgtc cagaattacg 16500

atcaggggta tgttgggttt cacgtctggc ctccggacca gcctccgctg gtccgattga 16560
acgcgcggat tctttatcac tgataagttg gtggacatat tatgtttatc agtgataaag 16620
tgtcaagcat gacaaagttg cagccgaata cagtgatccg tgccgccctg gacctgttga 16680
acgaggtcgg cgtagacggt ctgacgacac gcaaactggc ggaacggttg ggggttcagc 16740
agccggcgct ttactggcac ttcaggaaca agcgggcgct gctcgacgca ctggccgaag 16800
ccatgctggc ggagaatcat acgcattcgg tgccgagagc cgacgacgac tggcgctcat 16860
ttctgatcgg gaatgccgc agcttcaggc aggcgctgct cgcctaccgc gatggcgcg 16920
gcatccatgc cggcacgca cggggcgac cgcatatgga aacggccgac gcgcagcttc 16980
gcttcctctg cgaggcgggt ttttcggccg gggacgccgt caatgcgctg atgacaatca 17040
gctacttcac tgttggggcc gtgcttgagg agcaggccgg cgacagcgat gccggcgagc 17100
gcggcgccac cgttgaacag gctccgctct cgccgctgtt gcgggcccgc atagacgcct 17160
tcgacgaagc cgggtccggc gcagcgctcg agcagggact cgcggtgatt gtcgatggat 17220
tggcgaaaag gaggctcgtt gtcaggaacg ttgaaggacc gagaaagggt gacgattgat 17280
caggaccgct gccggagcgc aaccactca ctacagcaga gccatgtaga caacatcccc 17340
tcccccttc caccgctca gacgccgta gcagcccgt acgggctttt tcatgccctg 17400
ccctagcgtc caagctcac ggccgcgctc ggctctctg gcggccttct ggcgctcttc 17460
cgcttcctcg ctactgact cgtgcgctc ggctgctcgg ctgcggcgag cggtatcagc 17520
tcactcaaag gcggaatac gggtatccac agaatcaggg gataacgcag gaaagaacat 17580
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcggtttt 17640
ccataggctc cggccccctg acgagcatca caaaaatcga cgctcaagtc agagggtggc 17700
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 17760

tctgtttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 17820
ggcgctttttc cgctgcataa ccctgcttcg gggtcattat agcgattttt tcggtatatc 17880
catcctttttt cgcacgatat acaggattttt gccaaaggggt tcgtgtagac tttccttggt 17940
gtatccaacg gcgtcagccg ggcaggatag gtgaagtagg cccacccgcg agcgggtggt 18000
ccttcttcac tgtcccttat tcgcacctgg cgggtgtcaa cgggaatcct gctctgcgag 18060
gctggccggc taccgccggc gtaacagatg agggcaagcg gatggctgat gaaaccaagc 18120
caaccaggaa gggcagccca cctatcaagg tgtactgcct tccagacgaa cgaagagcga 18180
ttgaggaaaa ggcggcgggc gccggcatga gcctgtcggc ctacctgctg gccgtcggcc 18240
agggtacaa aatcacgggc gtcgtggact atgagcacgt ccgcgagctg gcccgcata 18300
atggcgacct gggccgctg ggcggcctgc tgaaactctg gctcaccgac gaccgcgca 18360
cggcgcggtt cggatgatgc acgatcctcg ccctgctggc gaagatcgaa gagaagcagg 18420
acgagcttgg caaggtcatg atgggcgtgg tccgcccag ggcagagcca tgactttttt 18480
agccgctaaa acggccgggg ggtgcgcgtg attgccaaagc acgtcccat gcgctccatc 18540
aagaagagcg acttcgcgga gctgggtgaag tacatcacgc acgagcaagg caagaccgag 18600
cgccttttgc acgctca 18617

<210> 51

<211> 18333

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 51

ccgggctggt tgcctctgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag	60
aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga tacctcgcgg	120
aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac	180
ccggcgccgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc	240
cagcctcgca aatcggcgaa aacgcctgat ttacgcgag ttccacacag atgatgtgga	300
caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat	360
gaggggcgcg atccttgaca cttgaggggc agagtgtga cagatgaggg gcgcacctat	420
tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt	480
ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg	540
tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc	600
cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgcccc	660
tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc cattgccggg	720
atcggggcag taacgggatg ggcgatcagc ccgagcgcca cgcccgaag cattgacgtg	780

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840
ggtggcggcc tgcccttcac ttcggcgcgc ggggcattca cggacttcat ggcggggccg 900
gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960
ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020
acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080
acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140
agataatata tcttttatat agaagatatc gccgtatgta aggatttcag ggggcaaggc 1200
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260
ctaattgcttg aaaccagga caataacctt atagcttgta aattctatca taattgggta 1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1380
agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc 1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatac cactcaaaag 1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta 1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgccccgc 1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa atcgtgttga 1800
ggccaacgcc cataatgcgg gctgttgccc ggcacccaac gccattcatg gccatatcaa 1860
tgattttctg gtgcgtaccg ggttgagaag cggtgtaagt gaactgcagt tgccatgttt 1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca 1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc 2040

aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca 2100

aaatcggctc cgtcgatact atgttatagc ccaactttga aaacaacttt gaaaaagctg 2160

ttttctggta ttttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat 2220

aattagcttc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc 2280

taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga 2340

tacggaagga atgtctcctg ctaaggtata taagctgggtg ggagaaaatg aaaacctata 2400

tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga 2460

catgatgcta tggctggaag gaaagctgcc tgttccaaag gtccctgcact ttgaacggca 2520

tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagagta 2580

tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctctt 2640

tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagccga 2700

attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaaga 2760

cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaaga 2820

ggaacttgtc ttttccacg gcgacctggg agacagcaac atctttgtga aagatggcaa 2880

agtaagtggc ttattgatc ttgggagaag cggcagggcg gacaagtggc atgacattgc 2940

cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctatTTTT 3000

tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120

tcttccgcac caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180

ggtcgctggc attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240

cggctctacg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300

gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgaag 3360
gagggtgaat gaatcggacg tttagaccga aggcatacag gcaagaactg atcgacgcgg 3420
ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgatcatgctg gcgccccgcg 3480
aaaccttcca gtccgtcggc tcgatgggtcc agcaagctac ggccaagatc gagcgcgaca 3540
gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600
gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660
tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720
agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780
tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840
ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960
acgaactggt gtggcagcag gtgttggagt acgcgaagcg caccctatc ggcgagccga 4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080
acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggccccgac 4320
ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380
gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500
tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca gcagccagcg 4560

ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc 4620
gctcgggacg cacggcgcg ctcacgaact gccgataaac agaggattaa aattgacaat 4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgcccta 4860
catcgacggc gagatcattg ggctgtcggc cttcaaacag gaggacggcc ccaaggacgc 4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 4980
cggatatgctg ctgcgggctg tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040
tccaacggga atctggtgga tgcgcattct catcctcggc gcacttaata tttcgctatt 5100
ctggagcttg ttgtttatct cggctctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160
cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220
attgatggcg gtcctggggg ctatttgcgg aactgcgggc gtggcgctgt tgggtgttgac 5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400
cgcttgga caatggcgagg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460
gccaatcccc atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg 5520
agcgggttta acctacttcc tttggttcgg ggggatctcg cgactcgaac ctacagttgt 5580
ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga tgcatacggc 5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700
tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcgggttta 5760
tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820

cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgga gatcatccgt 5940

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000

tctgccgect tacaacggct ctcccgtga cgccgtcccg gactgatggg ctgcctgtat 6060

cgagtgggtga ttttgtgccg agctgccggc cggggagctg ttggctggct ggtggcagga 6120

tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180

taatgtactg ggggtggttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240

accgcctggc cctgagagag ttgcagcaag cgggtccacgc tggtttgccc cagcaggcga 6300

aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360

ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480

cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540

ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660

cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6720

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780

ctcggtaacc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

cgaggagcct cgtcctgtca caactacca catggagtac gataagggcc agttccgcca 6900

gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcac tgtacttcaa 6960

gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctgggt 7020

aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgctccctt taaggctgcc 7080

aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140

aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200

cgggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260

tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320

tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380

tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcg catgaggttt 7440

tcgaaataca tccgatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7500

gcactattga tcatccgata gctctgcaaa gggcggttga caatgcaagt caaacgttgc 7560

tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620

tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740

atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800

cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatac 7920

tgtccagatc atggttgacc ggtgcctgga tcttctata gaatcatcct tattcgttga 7980

cctagctgat tctggagtga cccagagggg catgacttga gcctaaaatac cgccgctcc 8040

accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100

tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160

acagctctgg cggctctgag gtgcagtgga tgattattaa tccgggaccg gccgcccctc 8220

cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaattctatt gcatcatcgg 8280

agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340
gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400
atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtaccggc 8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520
ctgctttgcc cgggtgtatga aaccggaaaag gccgctcagg agctggccag cggcgcagac 8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640
tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca 8700
aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760
cttttctttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820
tatttcpcct aagtaagtac ttgctacat ccatactcca tccttcccat cccttattcc 8880
tttgaacctt tcagttcgag ctttccact tcatcgcagc ttgactaaca gctacccgc 8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgagaagt ctcgctgttt 9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120
ctacaaagat cgttatgttt atcggcactt tgcacggcc gcgctcccga ttccggaagt 9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgtc gttctgcagc cggtcgcgga 9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gccatttcg 9360
accgcaagga atcgggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc 9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgctccg tcgcgcaggc 9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc 9540

ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg 9600

gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc 9660

gtggttggt tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc 9720

aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac tctatcagag 9780

cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt 9840

ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg 9900

gaccgatggc tgtgtagaag tactcgccga tagtggaac cgacgcccc gactcgtcc 9960

gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa 10020

tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgcgatgtc agctccggag 10080

ttgagacaaa tgggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga 10140

gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta 10200

cctcttcag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc 10260

cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac 10320

gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc 10380

tccacgcgac tatatatattg tctctaattg tactttgaca tgctcctctt ctttactctg 10440

atagcttgac tatgaaaatt ccgtcaccag cncctgggtt cgcaaagata attgcatgtt 10500

tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat 10560

canttcctac taagatggta tacaatagta accatgcatg gttgcctagt gaatgctccg 10620

taacacccaa tacgccggcc gaaactttt tacaactctc ctatgagtcg ttaccaga 10680

atgcacaggt acacttggtt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740

gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaa 10800

tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860
atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920
aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaaagt ttttttttat 10980
acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040
atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100
tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat tttactaatt 11160
cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220
ttgattacac tctaaatctc gatgacccga caaaaagag cgaacaaaga aataatcttg 11280
tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340
tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt attttatctc 11400
tatgttgtgg atttggaatg ccctgatcgt tttcgttacc gtgattggca tgggaagtga 11460
tgctgcactg gcacacaaat acatcatgca cggctggggg tggggatggc atctttcaca 11520
tcatgaaccg cgtaaagggtg cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580
attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640
cgcaggtatg acggcgtatg gattactcta ttttatggtg cacgacgggc tgggtgatca 11700
acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760
ccgtatgcat cacgccgtca ggggcaaaga aggttgtggt tcttttggct tcctctatgc 11820
gccgcccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcgggcgc 11880
tgccagagat gcgcagggcg gggaggatga gcccgcatcc gggaagtaag ggctgacca 11940
gaggcggcca gcagcagcgt taatttttcg ggcgtggctg ttgactgccg ctgatcccaa 12000
agcttgagat taaaatagat aaggaaaaga aagtgaaaag aaattcggaa gcatggcaca 12060

ttcttctttt tataaatata tgcttgactt tctttttcca tcgatatgat atatgcatat 12120
gatagatata caagcaatct tcttcaagga gtttgaaatt ttgtcctcca ggagcaaaaa 12180
aaagtttttt tttatacatg tttgtacaca agaatagtta ccaatttgct ttggtcttac 12240
gtgctgcaag tttatatcgt tttcaatttc tttgtcttta cattttcttt gtcctttatc 12300
tttcctcatt tagtctttgg gagaattagg aaaagggagc ggaaaggtaa gaaatgcttg 12360
cgtattttac taattcggca aacatccaat ttggcaaaca gcagcctgtg caacgctctc 12420
gagatgacag tatctttgat tacactctaa atctcgatga cccgaccaa aagagcgaac 12480
aaagaaataa tcttgtgcat tcgaatatga tggaagattt tttccccctt attctaaatg 12540
ttgacatagc gtgtatgtta tataaacaaa aagaaattgt acaaactttc ttttcttctc 12600
tttttatttt atctctatga tccagttaga acaaccactc agtcatcaag caaaactgac 12660
tccagtactg agaagtaa atctcagttta ggggcttttc attgctattg tcattgttag 12720
cgcatgggtc attagcctga gtttattact ttcccttgac atctcaaagc taaaattttg 12780
gatgttattg cctgtttatac tatggcaaac atttttatat acgggattat ttattacatc 12840
tcatgatgcc atgcatggcg tagtatttcc ccaaaacacc aagattaatc atttgattgg 12900
aacattgacc ctatcccttt atggtctttt accatatcaa aaactattga aaaaacattg 12960
gttacaccac cacaatccag caagctcaat agaccggat tttcacaatg gtaaacacca 13020
aagtttcttt gcttgggtatt ttcattttat gaaaggttac tggagttggg ggcaaataat 13080
tgcggtgact attatttata actttgctaa atacatactc catatcccaa gtgataatct 13140
aacttacttt tgggtgctac cctcgctttt aagttcatta caattattct attttggtac 13200
ttttttaccc catagtgaac caataggggg ttatgttcag cctcattgtg cccaaacaat 13260
tagccgtcct atttgggtgt catttatcac gtgctatcat tttggctacc acgaggaaca 13320

tcacgaatat cctcatatctt cttggtggca gttaccagaa atttacaag caaaatagaa 13380
gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc 13440
cacacaacat acgagccgga agcataaagt gtaaagcctg gggcgcctaa tgagtgaagt 13500
aactcacatt aattgcgttg cgctcactgc cgcgtttcca gtcgggaaac ctgtcgtgcc 13560
agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggccaaaga 13620
caaaagggcg acattcaacc gattgagggg ggaaggtaa atattgacgg aaattattca 13680
ttaaagggtga attatcacctg tcaccgactt gagccatttg ggaattagag ccagcaaaat 13740
caccagtagc accattacca ttagcaaggc cggaaacgtc accaatgaaa ccatcgatag 13800
cagcaccgta atcagtagcg acagaatcaa gtttgccttt agcgtcagac tgtagcgcgt 13860
tttcatcggc attttcggtc atagccccct tattagcgtt tgccatcttt tcataatcaa 13920
aatcaccgga accagagcca ccaccggaac cgcctccctc agagccgcca ccctcagaac 13980
cgccaccctc agagccacca ccctcagagc cgccaccaga accaccacca gagccgccgc 14040
cagcattgac aggaggcccg atctagtaac atagatgaca ccgcgcgcga taatttatcc 14100
tagtttgcgc gctatatctt gttttctatc gcgtattaaa tgtataattg cgggactcta 14160
atcataaaaa cccatctcat aaataacgtc atgcattaca tgttaattat tacatgctta 14220
acgtaattca acagaaatta tatgataatc atcgcaagac cggcaacagg attcaatctt 14280
aagaaacttt attgccaaat gtttgaacga tcggggatca tccgggtctg tggcgggaac 14340
tccacgaaaa tatccgaacg cagcaagata tcgcggtgca tctcggctct gcctgggcag 14400
tcgccgcga cgccgttgat gtggacgccg ggcccgatca tattgtcgtc caggatcgtg 14460
gcgttggtgt tgtcggccgt tgctgtcgta atgatatcgg caccttcgac cgctgttcc 14520

gcagagatcc cgtgggcgaa gaactccagc atgagatccc cgcgctggag gatcatccag 14580
ccggcgcccc ggaaaacgat tccgaagccc aacctttcat agaaggcggc ggtggaatcg 14640
aaatctcgtg atggcaggtt gggcgtcgtt tggtcgggtca tttcgaaccc cagagtcccc 14700
ctcagaagaa ctctgaaga aggcgataga aggcgatgag ctgcgaatcg ggagcggcga 14760
taccgtaaag cagaggaag cggtcagccc attcgccgcc aagctcttca gcaatatcac 14820
gggtagccaa cgctatgtcc tgatagcggc ccgccacacc cagccggcca cagtcgatga 14880
atccagaaaa ggggccattt tccaccatga tattcggcaa gcaggcatcg ccatgggtca 14940
cgacgagatc atcgccgtcg ggcagtcgag ccttgagcct ggcgaacagt tcggctggcg 15000
cgagcccctg atgctcttcg tccagatcat cctgatcgac aagaccggct tccatccgag 15060
tacgtgctcg ctgatgaga tgtttcgtt ggtggtcgaa tgggcaggta gccggatcaa 15120
gcgtatgcag ccgccgcatt gcatcagcca tgatggatac tttctcggca ggagcaaggc 15180
gagatgacag gagatcctgc cccggcactt cgcccaatag cagccagtcc cttcccgtt 15240
cagtgacaac gtcgagcaca gctgcgcaag gaacgcccggt cgtggccagc cacgatagcc 15300
gcgctgcctc gtctgcagt tcattcaggg caccggacag gtcgggtcttg aaaaaagaa 15360
ccgggcgccc ctgcgtgac agccggaaca cggcggcatc agagcagccg attgtctgtt 15420
gtgcccagtc atagccgaat agcctctcca cccaagcggc cggagaacct gcgtgcaatc 15480
catcttggtc aatcatgca aacgatccag atccgggtgca gattatttgg attgagagtg 15540
aatatgagac tctaattgga taccgagggg aatttatgga acgtcagtgg agcatttttg 15600
acaagaaata tttgctagct gatagtgacc ttaggcgact tttgaacgcg caataatggt 15660
ttctgacgta tgtgcttagc tcattaaact ccagaaaccc gcggctgagt ggctccttca 15720
acgttgcggt tctgtcagtt ccaaacgtaa aacggcttgt cccgcgtcat cggcgggggt 15780

cataacgtga ctcccttaat tctccgctea tgatcagatt gtcgtttccc gccttcagtt 15840
taaactatca gtgtttgaca ggatatattg gcgggtaaac ctaagagaaa agagcgttta 15900
ttagaataat cggatattta aaagggcgtg aaaaggttta tccgttcgtc catttgtatg 15960
tgcatgccaa ccacagggtt cccagatctt ggcgccggcc agcgagacga gcaagattgg 16020
ccgccgccc aaacgatccg acagcgcgcc cagcacaggt gcgcaggcaa attgcaccaa 16080
cgcatacagc gccagcagaa tgccatagtg ggcggtgacg tcgttcgagt gaaccagatc 16140
gcgcaggagg cccggcagca ccggcataat caggccgatg ccgacagcgt cgagcgcgac 16200
agtgtcaga attacgatca ggggtatgtt gggtttcacg tctggcctcc ggaccagcct 16260
ccgctggctc gattgaacgc gcggattctt tatcactgat aagttgggtgg acatattatg 16320
tttatcagtg ataaagtgtc aagcatgaca aagttgcagc cgaatacagt gatccgtgcc 16380
gccctggacc tgttgaacga ggtcggcgta gacggtctga cgacacgcaa actggcgga 16440
cggttggggg ttcagcagcc ggcgctttac tggcacttca ggaacaagcg ggcgctgctc 16500
gacgcactgg ccgaagccat gctggcggag aatcatacgc attcggtgcc gagagccgac 16560
gacgactggc gctcatttct gatcgggaat gcccgagct tcaggcaggc gctgctcgcc 16620
taccgcgatg gcgcgcgcac ccatgccggc acgcgaccgg gcgcaccgca gatggaaacg 16680
gccgacgcgc agcttcgctt cctctgcgag gcgggttttt cggccgggga cgccgtcaat 16740
gcgctgatga caatcagcta cttcactgtt ggggccgtgc ttgaggagca ggccggcgac 16800
agcgatgccg gcgagcgcg cggcaccgtt gaacaggctc cgctctcgcc gctgttgcg 16860
gccgcgatag acgccttcga cgaagccggt ccggacgcag cgctcgagca gggactcgcg 16920
gtgattgtcg atggattggc gaaaaggagg ctcgttgtca ggaacgttga aggaccgaga 16980
aagggtgacg attgatcagg accgctgccg gagcgcaacc cactcactac agcagagcca 17040

tgtagacaac atccccctccc cctttccacc ggcgcagacg cccgtagcag cccgctacgg 17100
gctttttcat gccctgccct agcgtccaag cctcacggcc gcgctcggcc tctctggcgg 17160
ccttctggcg ctcttcogct tcctcgtca ctgactcgt gcgctcggtc gttcggctgc 17220
ggcgagcgggt atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 17280
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 17340
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 17400
caagtacagag gtggcgaaac ccgacaggac tataaagata ccaggcggtt cccctggaa 17460
gtccctcgt gcgctctct gtccgaccc tgccgcttac cggatacctg tccgccttcc 17520
tcccttcggg aagcgtggcg cttttccgct gcataaccct gcttcgggggt cattatagcg 17580
attttttcgg tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt 17640
gtagactttc ctgggtgtat ccaacggcgt cagccgggca ggataggtga agtaggcccc 17700
cccgcgagcg ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg 17760
aatcctgctc tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg 17820
gctgatgaaa ccaagccaac caggaagggc agcccaccta tcaaggtgta ctgccttcca 17880
gacgaacgaa gagcgattga ggaaaaggcg gcggcggccg gcatgagcct gtcggcctac 17940
ctgctggccg tcggccaggg ctacaaaatc acgggcgtcg tggactatga gcacgtccgc 18000
gagctggccc gcatcaatgg cgacctgggc cgctggggcg gcctgctgaa actctggctc 18060
accgacgacc cgcgcacggc gcggttcggt gatgccacga tcctcgccct gctggcgaag 18120
atcgaagaga agcaggacga gcttggcaag gtcatgatgg gcgtgggtccg cccgagggca 18180
gagccatgac ttttttagcc gctaaaacgg ccgggggggtg cgcgtgattg ccaagcacgt 18240
cccatgcgc tccatcaaga agagcgactt cgcgagctg gtgaagtaca tcaccgacga 18300

gcaaggcaag accgagcgcc ttgcgacgc tca

18333

<210> 52

<211> 17

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (9)..(9)

<223> n is a, c, g, or t

<400> 52

gcngarggna thtggtta

17

<210> 53

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (6)..(6)

<223> n is a, c, g, or t

<400> 53

tcngcnagra adatrtrrtg

20

<210> 54

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 54

aagtgacacc ggttacacgc ttgtctt

27

<210> 55

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 55

gcttatcacc atctgttacc tccttgc

27

<210> 56

<211> 32

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 56

agagagggat ccttaaatgc gaatatcggtt gc

32

<210> 57
<211> 32
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 57
agagagggat ccatgtctga tcaaaagaag ca

32

<210> 58
<211> 37
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 58
actttattgg atccttaa at gcgaatatcg ttgctgc

37

<210> 59
<211> 38
<212> DNA
<213> Artificial

<220>
<223> Primer

<400> 59
gttccaattg gccacatgaa gagtaagaca ggaaacag

38

<210> 60
<211> 38
<212> DNA
<213> Artificial

<220>

<223> Primer

<400> 60

cctgtctttac tcttcatgtg gccaatggga accaacac

38

<210> 61

<211> 38

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 61

ctatttttaat catatgtctg atcaaaagaa gcatattg

38

<210> 62

<211> 16103

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 62

gatctttcga cactgaaata cgtcgcgcct gctccgcttg gaagcggcga ggagcctcgt	60
cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc	120
cagttcatgg gcgttgccat gatggccgctc atgcatctgt acttcaagta caccaacgct	180
cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct	240
tccggaagcc agcgactggg gacctccagc gtccctttaa ggctgccaac agctttctca	300
gccagggccca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg	360
gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg	420
atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca	480
tttggatttg ctttccaggc tgagactcta gcttgagca tagagggtcc tttggctttc	540
aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga	600
gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc	660
ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca	720
tccgatagct ctgcaaaggc cgttgcacaa tgcaagtcaa acgttgctag cagttccagg	780
tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct	840
caccacaaaa gtcagacggc gtaacacaaa gtcacacaac acaagctgta aggatttcgg	900
cacggctacg gaagacggag aagccacctt cagtggactc gaggaccatt taattctatt	960
tgtgtttgat cgagacctaa tacagcccct acaacgacca tcaaagtcgt atagctacca	1020
gtgaggaagt ggactcaaat cgacttcagc aacatctcct ggataaactt taagcctaaa	1080
ctatacagaa taagataggt ggagagctta taccgagctc ccaaattctgt ccagatcatg	1140
gttgaccggt gcctggatct tcctatagaa tcaccttat tcgttgacct agctgattct	1200

ggagtgaccc agagggtcat gacttgagcc taaaatccgc cgcctccacc attttagaaa 1260
aatgtgacg aactcgtgag ctctgtacag tgaccggtga ctctttctgg catgcggaga 1320
gacggacgga cgcagagaga agggctgagt aataagccac tggccagaca gctctggcgg 1380
ctctgaggtg cagtggatga ttattaatcc gggaccggcc gccctccgc cccgaagtgg 1440
aaaggctggt gtgcccctcg ttgaccaaga atctattgca tcatcggaga atatggagct 1500
tcatcgaatc accggcagta agcgaaggag aatgtgaagc caggggtgta tagccgtcgg 1560
cgaaatagca tgccattaac ctaggtacag aagtccaatt gcttccgac tggtaaaaga 1620
ttcacgagat agtaccttct ccgaagtagg tagagcgagt acccggcgcg taagctccct 1680
aattggccca tccggcatct gtagggcgtc caaatatcgt gcctctcctg ctttgcccgg 1740
tgtatgaaac cggaaaggcc gctcaggagc tggccagcgg cgcagaccgg gaacacaagc 1800
tggcagtcga cccatccggt gctctgcact cgacctgctg aggtccctca gtccctggta 1860
ggcagctttg ccccgctctgt ccgcccgggtg tgtcggcggg gttgacaagg tcgttgcgtc 1920
agtccaacat ttgttgccat attttctcgc tctcccacc agctgctctt ttcttttctc 1980
tttcttttcc catcttcagt atattcatct tcccatccaa gaacctttat ttcccctaag 2040
taagtacttt gctacatcca tactccatcc ttcccatccc ttattccttt gaacctttca 2100
gttcgagctt tcccacttca tcgcagcttg actaacagct acccgccttg agcagacatc 2160
accatgcctg aactcacgcg gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 2220
gtctccgacc tgatgcagct ctccggagggc gaagaatctc gtgctttcag cttcgatgta 2280
ggagggcggtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 2340
tatgtttatc ggcactttgc atcgcccgcg ctcccgattc cggaagtgtc tgacattggg 2400
gaattcagcg agagcctgac ctattgcac tccgcgctg cacagggtgt cacgttgcaa 2460

gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg 2520
atcgctgegg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc 2580
ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac 2640
tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg 2700
atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga tttcggctcc 2760
aacaatgtcc tgacggacaa tggccgcata acagcggtea ttgactggag cgaggcgatg 2820
ttcgggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt 2880
atggagcagc agacgcgcta cttcgagcgg aggcattccg agcttgcagg atcgccgcgg 2940
ctccggggcg atatgtccg cattggctct gaccaactct atcagagctt ggttgacggc 3000
aatttcgatg atgcagcttg ggcgcagggc cgatgcgacg caatcgtccg atccggagcc 3060
gggactgtcg ggcgtacaca aatcgcccg cagaagcgcgg ccgtctggac cgatggctgt 3120
gtagaagtac tcgccgatag tggaaaccga cgcgccagca ctcgtccgag ggcaaaggaa 3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct 3240
tgacgaatct ggatataaga tcgttgggtg cgatgtcagc tccggagttg agacaaatgg 3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagtg 3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata 3420
cagctcatct gcaatgcatt aatgcattga ctgcaaccta gtaacgcctt ncaggctccg 3480
gcgaagagaa gaatagctta gcagagctat tttcattttc gggagacgag atcaagcaga 3540
tcaacggtcg tcaagagacc tacgagactg aggaatccgc tcttggctcc acgcgactat 3600
atatttgtct ctaattgtac tttgacatgc tcctcttctt tactctgata gcttgactat 3660
gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720

ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780
gatggtatac aatagtaacc atgcatgggt gcctagtga tgctccgtaa cacccaatac 3840
gccggccgaa acttttttac aactctccta tgagtcgttt acccagaatg cacaggatca 3900
cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960
ctccacatct ccaactcgacc tgcaggcatg caagcttgag tctatcgctt ccaaaaagta 4020
cggtgctgaa ttcagatata aatcgctgtg tgctaaaatt aacactgtcg ataaagacaa 4080
gcgtgtaacc ggtgtcactt tggaaagcgg agaagtcatt gaagccgatg cagtcgtatg 4140
taatgcggat cttgtttatg cttatcacca tctgttacct ccttgcaatt ggacaaagaa 4200
gacattagcc tcaaagaaac tcaacttcac atctatttcg ttttattggg ccatgtcaac 4260
aaaggtgcct caattagacg tacacaatat cttcttggct gaagcctaca aggaaagttt 4320
tgatgagatt ttcaacgact tcggtttgcc ctctgaagct tggcgtaatc atgggtcatag 4380
ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc 4440
ataaagtgtg aagcctgggg tgcctaata gtagagtaac tcacattaat tgcgttgctg 4500
tcaactgccc ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa 4560
cgcgcgggga gaggcgggtt gcgtattggg ccaaagacaa aaggcgacaa ttcaaccgat 4620
tgaggggagg aaggtaaata ttgacggaaa ttattcatta aaggtgaatt atcacgtca 4680
ccgacttgag ccatttggga attagagcca gcaaatcac cagtagcacc attaccatta 4740
gcaaggccgg aaacgtcacc aatgaaacca tcgatagcag caccgtaatc agtagcgaca 4800
gaatcaagtt tgcccttagc gtcagactgt agcgcggttt catcggcatt ttcggtcata 4860
gcccccttat tagcggttgc catcttttca taatcaaaat caccggaacc agagccacca 4920
ccggaaccgc ctccctcaga gccgccacc tcagaaccgc caccctcaga gccaccacc 4980

tcagagccgc caccagaacc accaccagag ccgccgccag cattgacagg aggccccgac 5040

tagtaacata gatgacaccg cgcgcgataa tttatcctag tttgcgcgct atattttgtt 5100

ttctatcgcg tattaatatgt ataattgcgg gactctaata ataaaaaccc atctcataaa 5160

taacgtcatg cattacatgt taattattac atgcttaacg taattcaaca gaaattatat 5220

gataatcatc gcaagaccg caacaggatt caatcttaag aaactttatt gccaaatgtt 5280

tgaacgatcg gggatcatcc gggctctgtg cggaactcc acgaaaatat ccgaacgcag 5340

caagatatcg cggatcatct cggctcttgc tgggcagtcg ccgccgacgc cgttgatgtg 5400

gacgccgggc ccgatcatat tgctgctcag gatcgtggcg ttgtgcttgt cggccgttgc 5460

tgctgtaatg atatcggcac ctctgaccgc ctgttccgca gagatcccgt gggcgaagaa 5520

ctccagcatg agatccccgc gctggaggat catccagccg gcgtcccgga aaacgattcc 5580

gaagcccaac ctttcataga aggcggcggt ggaatcgaaa tctcgtgatg gcaggttggg 5640

cgctcgttgg tgggtcattt cgaaccccag agtcccgtc agaagaactc gtcaagaagg 5700

cgatagaagg cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg 5760

tcagcccatt cgcgcgaag ctcttcagca atatcacggg tagccaacgc tatgtcctga 5820

tagcgggtccg ccacaccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc 5880

accatgatat tcggcaagca ggcacgcga tgggtcacga cgagatcatc gccgtcgggc 5940

atgcgcgcct tgagcctggc gaacagttcg gctggcgca gccctgatg ctcttcgtcc 6000

agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt 6060

ttcgcttggg ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca 6120

tcagccatga tggatacttt ctcggcagga gcaaggtag atgacaggag atcctgcccc 6180

ggcacttcgc ccaatagcag ccagtcctt cccgcttcag tgacaacgct gagcacagct 6240

gcgcaaggaa cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca 6300

ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc 6360

cggaacacgg cggcatcaga gcagccgatt gtctgttgtg cccagtcata gccgaatagc 6420

ctctccaccc aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaaac 6480

gatccagatc cggatgcagat tatttggtt gagagtgaat atgagactct aattggatac 6540

cgaggggaat ttatggaacg tcagtggagc atttttgaca agaaatattt gctagctgat 6600

agtgcctta ggcgactttt gaacgcgcaa taatggtttc tgacgtatgt gcttagctca 6660

ttaaactcca gaaacccgcg gctgagtggc tccttcaacg ttgcggttct gtcagttcca 6720

aacgtaaaac ggcttgtccc gcgtcatcgg cgggggtcat aacgtgactc ccttaattct 6780

ccgctcatga tcagattgtc gtttcccgcc ttcagtttaa actatcagtg tttgacagga 6840

tatattggcg ggtaaacta agagaaaaga gcgtttatta gaataatcgg atatttaaaa 6900

gggcgtgaaa aggtttatcc gttcgtccat ttgtatgtgc atgccaacca cagggttccc 6960

cagatctggc gccggccagc gagacgagca agattggccg ccgccgaaa cgatccgaca 7020

gcgcgcccag cacagggtgc caggcaaatt gcaccaacgc atacagcgcc agcagaatgc 7080

catagtgggc ggtgacgtc ttcgagtga ccagatcgcg caggaggccc ggcagcaccg 7140

gcataatcag gccgatgccg acagcgtcga gcgcgacagt gtcagaatt acgatcaggg 7200

gtatgttggg tttcacgtct ggcctccgga ccagcctccg ctggtccgat tgaacgcgcg 7260

gattctttat cactgataag ttggtggaca tattatgttt atcagtgata aagtgtcaag 7320

catgacaaaag ttgcagccga atacagtgat ccgtgccgcc ctggacctgt tgaacgaggt 7380

cggcgtagac ggtctgacga cagcaaaact ggcggaacgg ttgggggttc agcagccggc 7440

gctttactgg cacttcagga acaagcgggc gctgctcgac gcactggdcg aagccatgct 7500

ggcggagaat catacgatt cgggtgccgag agccgacgac gactggcgct cattttctgat 7560
cggaatgcc cgcagcttca ggcaggcgct gctcgcctac cgcgatggcg cgcgcatcca 7620
tgccggcacg cgaccgggag caccgcagat ggaaacggcc gacgcgacgc ttgccttctt 7680
ctgcgaggcg ggttttttcgg ccggggacgc cgtcaatgcg ctgatgacaa tcagctactt 7740
cactgttggg gccgtgcttg aggagcaggc cggcgacagc gatgccggcg agcgcggcg 7800
caccgttgaa caggctccgc tctcgccgct gttgcgggccc gcgatagacg ccttcgacga 7860
agccgggtccg gacgcagcgt tcgagcaggg actcgcgggtg attgtcgatg gattggcgaa 7920
aaggaggctc gttgtcagga acgttgaagg accgagaaaag ggtgacgatt gatcaggacc 7980
gctgccggag cgcaaccac tctactacgc agagccatgt agacaacatc ccctccccct 8040
ttccaccgag tcagacgccc gtagcagccc gctacgggct ttttcatgcc ctgccctagc 8100
gtccaagcct cacggccgag ctcggcctct ctggcgccct tctggcgctc ttccgcttcc 8160
tcgctcactg actcgtgcg ctcggtcggt cggctgcggc gagcgggtatc agctcactca 8220
aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca 8280
aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcggt tttccatagg 8340
ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagagggtg gcgaaacccg 8400
acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctggt 8460
ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt 8520
ttccgctgca taaccctgct tcggggtcct tatagcgatt ttttcgggtat atccatcctt 8580
tttcgcacga tatacaggat ttgcccagg ggttcgtgta gactttcctt ggtgtatcca 8640
acggcgtcag ccgggcagga taggtgaagt agggccaccc gcgagcgggt gttccttctt 8700
cactgtccct tattcgcacc tggcggtgct caacgggaat cctgctctgc gaggctggcc 8760

ggctaccgcc ggcgtaacag atgagggcaa gcggatggct gatgaaacca agccaaccag 8820

gaagggcagc ccacctatca aggtgtactg ccttccagac gaacgaagag cgattgagga 8880

aaaggcggcg gcggccggca tgagcctgtc ggcctacctg ctggccgctcg gccagggcta 8940

caaaatcacg ggcgtcgtgg actatgagca cgtccgagag ctggcccgca tcaatggcga 9000

cctgggccgc ctgggcggcc tgctgaaact ctggctcacc gacgacccgc gcacggcgcg 9060

gttcggtgat gccacgatcc tcgccctgct ggcgaagatc gaagagaagc aggacgagct 9120

tggcaaggtc atgatggcg tggtccgccc gagggcagag ccatgacttt ttagccgct 9180

aaaacggccg gggggtgcgc gtgattgcca agcacgtccc catgcgctcc atcaagaaga 9240

gcgacttcgc ggagctggtg aagtacatca ccgacgagca aggcaagacc gagcgccttt 9300

gcgacgctca ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca 9360

aacgcgccag aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttggtga 9420

tacctcgcg aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg 9480

gccgactcac ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg 9540

tggagctggc cagcctcgca aatcggcgaa aacgcctgat tttacgagag tttccacag 9600

atgatgtgga caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact 9660

actgacagat gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg 9720

gcgcacctat tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt 9780

ttccgcccgt ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt 9840

ataaaccttg tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg 9900

ggtgcccccc cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg 9960

gctgcgcccc tcggccgcga acggcctcac ccaaaaaatg gcagcgctgg cagtccttgc 10020

cattgccggg atcggggcag taacgggatg ggcgatcagc ccgagcgca cgcccgggaag 10080

cattgacgtg ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg 10140

cggcggcctg ggtggcgcc tgcccttcac ttcggccgtc ggggcattca cggacttcat 10200

ggcggggccg gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct 10260

cgtgttcggg ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg 10320

aggtatgaaa acgagaattg gacctttaca gaattactct atgaagcgcc atatttataaa 10380

agctaccaag acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac 10440

aatactgata agataatata tcttttatat agaagatata gccgtatgta aggatttcag 10500

ggggcaaggc ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa 10560

cttgcattga ctaatgcttg aaaccagga caataacctt atagcttgta aattctatca 10620

taattgggta atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac 10680

tttgtcatgc agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag 10740

gtgctgcctc agattcaggt tatgccgtc aattcgtgc gtatatcgct tgctgattac 10800

gtgcagcttt cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac 10860

cacgtcaaag ggtgacagca ggctcataag acgcccagc gtcgcatag tgcgttcacc 10920

gaatacgtgc gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg 10980

gcgcgattta gcccgcacat agcccactg ttcgtccatt tccgcgcaga cgatgacgtc 11040

actgcccggc tgtatgcgcg aggttaccga ctgcggcctg agttttttta gtgacgtaaa 11100

atcgtgttga ggccaacgcc cataatgcgg gctgttgccc ggcattccaac gccattcatg 11160

gccatatcaa tgattttctg gtgcgtaccg ggttgagaag cgggtgaagt gaactgcagt 11220

tgccatgttt tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg 11280

ttacgcacca ccccgctcagt agctgaacag gagggacagc tgatagacac agaagccact 11340
ggagcacctc aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat 11400
tgtgggtttca aaatcggtc cgctgatact atgttatatcg ccaactttga aaacaacttt 11460
gaaaaagctg ttttctggta tttaagggtt tagaatgcaa ggaacagtga attggagttc 11520
gtcttggtat aattagottc ttgggggtatc tttaaatact gtagaaaaga ggaaggaaat 11580
aataaatggc taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct 11640
gcgtaaaaga tacggaagga atgtctcctg ctaagggtata taagctgggtg ggagaaaatg 11700
aaaacctata tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac 11760
gggaaaagga catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact 11820
ttgaacggca tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct 11880
cggaagagta tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgc 11940
tcaggctctt tcaactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc 12000
gcttagccga attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact 12060
gggaagaaga cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa 12120
agcccgaaga ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga 12180
aagatggcaa agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtgg 12240
atgacattgc cttctgcgtc cggctgatca gggaggatat cggggaagaa cagtatgtcg 12300
agctattttt tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt 12360
tactggatga attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag 12420
cgcaccgact tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat 12480
ttgggcaagg ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag 12540

gacggccaga cggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc 12600
aaggcaccag gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca 12660
atccccgaag gaggggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg 12720
atcgacgcgg ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt 12780
gcgccccgcg aaaccttcca gtccgtcggc tcgatgggtcc agcaagctac ggccaagatc 12840
gagcgcgaca gcgtgcaact ggctccccct gccctgcccc cgccatcggc cgccgtggag 12900
cgttcgcgtc gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg 12960
cgaggaacta tgacgaccaa gaagcgaaaa accgcccggc aggacctggc aaaacaggtc 13020
agcgaggcca agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag 13080
ctttccttgt tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg 13140
gcccgctctg ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac 13200
aaggtcattt tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg 13260
gccgacgatg acgaactggg gtggcagcag gtgttgaggt acgcgaagcg caccctatc 13320
ggcgagccga tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat 13380
ggccggtatt acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc 13440
ttcacgtccg accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc 13500
ctggaccgtg gcaagaaaac gtcccgttgc caggctctga tcgacgagga aatcgtcgtg 13560
ctgtttgctg gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg 13620
acggccccgac ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg 13680
gaaaccttcc gcctcatgtg cggatcggat tccaccgcg tgaagaagtg gcgcgagcag 13740
gtcggcgaag cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat 13800

gatgacctgg tgcattgcaa acgctagggc cttgtggggg cagttccggc tgggggttca 13860
gcagccagcg ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc 13920
gctcagtatc gctcgggacg cacggcgcg cttacgaact gccgataaac agaggattaa 13980
aattgacaat tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt 14040
tccgcgagat ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg 14100
agcacgagga gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat 14160
tcggcgcccta catcgacggc gagatcattg ggctgtcggg cttcaaacag gaggacggcc 14220
ccaaggacgc tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc 14280
gaggggtcgc cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg 14340
tccgacagat tccaacggga atctgggtga tgcgcatctt catcctcggc gcacttaata 14400
tttcgctatt ctggagcttg ttgtttattt cgggtctaccg cctgccgggc ggggtcgcgg 14460
cgacggtagg cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta 14520
gcccgatagc attgatggcg gtccctgggg ctatttgcgg aactgcgggc gtggcgctgt 14580
tggtgttgac accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg 14640
cggtttccat ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc 14700
tcacctttac cgcttgga caaggcgccg gaggacttct gctcgttcca gtagctttag 14760
tgtttgatcc gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg 14820
gcctgatcgg agcgggttta acctacttcc ttgggttcg ggggatctcg cgactcgaac 14880
ctacagttgt ttccttactg ggctttctca gcccagatc tggggtcgat cagccgggga 14940
tgcatcaggc cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg 15000
ataggggagt tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc 15060

agcggcttta tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt 15120
cacggttaag cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga 15180
tatttgatca caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga 15240
gatcatccgt gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac 15300
atgagcaaag tctgccgcct tacaacggct ctcccgtga cgccgtcccg gactgatggg 15360
ctgcctgtat cgagtgggtga ttttgtgccg agctgccggt cggggagctg ttggctggct 15420
gggtggcagga tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg 15480
cggacgtttt taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg 15540
attgcccttc accgcctggc cctgagagag ttgcagcaag cggtcacgc tggtttgccc 15600
cagcaggcga aaatcctgtt tgatgggtgg tccgaaatcg gcaaaatccc ttataaatca 15660
aaagaatagc ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta 15720
aagaacgtgg actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta 15780
cgtgaaccat cacccaaate aagttttttg gggtcgaggt gccgtaaagc actaaatcgg 15840
aacccataag ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga 15900
aaggaagggg agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg 15960
gcgatcgggtg cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag 16020
gcgattaagt tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag 16080
tgaattcgag ctcggtaccc ggg 16103

<210> 63

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 63

ggcgtacttg aaggaaccct taccg

25

<210> 64

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 64

attgatgctc ccggtcaccg tgatt

25

<210> 65

<211> 500

<212> DNA

<213> Blakeslea trispora

<400> 65

aatctataca atgctccata gactcacatt gatattgtcg aagatttcga tgctgactta 60

gtagagcaac tacaaaagtt agcagagaag catgatttct taatctttga agaccgcaag 120

tttgcagata tcggtatgtg aattctatct attttttttc tgatgtgtgc atggatgact 180

catgatcata ttcttaggta atactgtcaa gcatcaatat ggcaagggcg ttacaagat 240

tgcttcttgg tctcatatta ctaatgctca cacagttcct ggagaaggta ttatcaaggg 300

acttgccgaa gtcggcctcc ctcttggtcg tggcttgctt ttgctagcag aaatgtcatc 360

tcaaggtgca ttaactaagg gtatttacac tgccgaatct gtcaatatgg ctgcgcgcaa 420

caaagatttc gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga 480

ttttgttgtc atgtcgctg

500

<210> 66

<211> 611

<212> DNA

<213> Blakeslea trispora

<400> 66

gagattaaaa tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt 60

ctttttataa atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag 120

atatacaagc aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaagt 180

ttttttttat acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct 240

gcaagtttat atcgttttca atttctttgt ctttacattt tctttgtcct ttatctttcc 300

tcatttagtc tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgat 360

tttactaatt cggcaaact ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat 420

gacagtatct ttgattacac tctaaatctc gatgaccgga ccaaaaagag cgaacaaaga 480

aataatcttg tgcattcgaa tatgatggaa gattttttcc cccttattct aaatggtgac 540

atagcgtgta tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctcttttt 600

attttatctc t 611

<210> 67

<211> 720

<212> DNA

<213> Blakeslea trispora

<400> 67

atgtcaatac tcacttatct ggaatttcat ctctactata cactacctgt ccttgcgga 60

ttgtgttggc tgctaaagcc gtttcactca cagcaagaca atctcaagta taaattttta 120

atgttgatgg cgcctctac cgcctcgatt tgggacaatt atatcgttta tcatcgcgct 180
tgggtggact gtcctacttg tgttgaggct gtcattggct atgtacctct agaagaatac 240
atgttcttta tcatcatgac tttaatgact gtcgcgttct caaactttgt tatgcgttgg 300
cacttgcata ctttctttat tagaccaac acttcttgga agcaaact attagtaggc 360
cttgtgcctg tttcagcttt attggcaatc acttatcatg cttggcactt gacactgcca 420
aataaacctt cattttatgg ttcatgcac ctttggtatg cttgtcctgt gttggctatt 480
ctttggctgg gtgctggcga atatatcttg cgtcgacctg tggctgtcct tttgtctatt 540
gttatcccta gtgtatacct atgttgggct gatatcgctg ctattagtgc tggcacatgg 600
catatttctc ttagaacaag cactggcaaa atggtagtac ccgatttacc tgtagaagaa 660
tgctgtttt ttactttgat caacacagtc ttggtttttg ctacctgtgc tatagaccgc 720

<210> 68

<211> 1089

<212> DNA

<213> *Blakeslea trispora*

<400> 68

ctgtacaaat catctgttca aaatcaaaac ctaaacaag ccatttcct tttccagcat 60
gtcaaagagc tagcatgggc cttctgtctt cctgaccaa tgctcaaca tgaattgttt 120
gatgatctta ctatcagctg ggatatttta cgtaaagcct caaagtcatt ctatactgca 180
tctgccgttt ttccaagtta tgtacgtcaa gacttgggtg ttctctatgc tttctgcaga 240
gctaccgatg acctgtgcga tgatgaatcc aaatctgttc aagaagaag agaccaatta 300
gatcttactc gacaatttgt tcgtgatctc tttagccaaa agaccagtgc gcctattgtg 360
attgattggg aattgtatca aaaccaactt cctgcttctt gtatatcagc ctttagagcc 420
tttactcgcc ttgcctatgt ccttgaagta gaccctgtag aagaactatt agatggttac 480

aaatgggatc ttgagcgctg tcctatcctt gatgaacaag acttggaggc atactctgct 540
tgtgtggcca gtagtgtggg tgaaatgtgc acacgtgtga ttcttgctca agaccaaag 600
gaaaatgatg cttggataat tgaccgtgca cgtgagatgg ggctgggtgct acaatacgtt 660
aacattgctc gagacattgt gactgatagc gagactctgg gtcgatgtta tctgcctcaa 720
caatggctta gaaaagaaga aacagaacaa atacagcaag gcaacgcccg tagcctaggt 780
gatcaaagac tggtgggctt gtctctgaag cttgtaggaa aggcagacgc tatcatggtg 840
agagctaaga agggcattga caagttgccg gcaaactgtc aaggcgggtgt acgagctgct 900
tgccaagtat atgctgcaat tggatctgta ctcaagcagc agaagacaac atatcctaca 960
agagctcatc taaaaggaag cgaacgtgcc aagattgctc tgttgagtgt atacaacctc 1020
tatcaatctg aagacaagcc tgtggctctc cgtcaagcta gaaagattaa gagttttttt 1080
gttgattag 1089

<210> 69

<211> 611

<212> DNA

<213> *Blakeslea trispora*

<400> 69

agagataaaa taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac 60
atacacgcta tgtcaacatt tagaataagg gggaaaaaat cttccatcat attogaatgc 120
acaagattat ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca 180
aagatactgt catctcgaga gcgttgca ca ggctgctgtt tgccaaattg gatgtttgcc 240
gaattagtaa aatacgcaag catttcttac ctttccgctc ctttttcccta attctcccaa 300
agactaaatg aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga 360

tataaacttg cagcacgtaa gaccaaagca aattggtaac tattcttggtg tacaacatg 420
tataaaaaaa aacttttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat 480
tgcttgata tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta 540
tttataaaaa gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct 600
attttaatct c 611

<210> 70

<211> 882

<212> DNA

<213> Haematococcus pluvialis

<400> 70

atgctgtcga agctgcagtc aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac 60
atcacgcggc ccaaagtctg cctgcatgct cagcgggtgct cgtagttcg gctgcgagtg 120
gcagcaccac agacagagga ggcgctggga accgtgcagg ctgccggcgc gggcgatgag 180
cacagcgccg atgtagcact ccagcagctt gaccgggcta tcgcagagcg tcgtgcccgg 240
cgcaaacggg agcagctgtc ataccaggct gccgccattg cagcatcaat tggcgtgtca 300
ggcattgcca tcttcgccac ctacctgaga ttgcatgc acatgaccgt gggcggcgca 360
gtgccatggg gtgaagtggc tggcactctc ctcttggtgg ttggtggcgc gctcggcatg 420
gagatgtatg cccgctatgc acacaaagcc atctggcatg agtcgcctct gggctggctg 480
ctgcacaaga gccaccacac acctcgact ggaccctttg aagccaacga cttgtttgca 540
atcatcaatg gactgcccgc catgctcctg tgtacctttg gcttctggct gcccaacgtc 600
ctggggggcg cctgcttttg agcgggggtg ggcatacgc tatacggcat ggcataatg 660
tttgtacacg atggcctggg gcacaggcgc tttcccaccg ggcccatcgc tggcctgccc 720
tacatgaagc gcctgacagt ggcccaccag ctacaccaca gcggcaagta cgggtggcgcg 780

ccctggggta tgttcttggg tccacaggag ctgcagcaca ttccagggtgc ggcggaggag 840

gtggagcgcac tggctctgga actggactgg tccaagcggg ag 882

<210> 71

<211> 528

<212> DNA

<213> *Erwinia uredovora*

<400> 71

atgttgtgga ttggaatgc cctgatcgtt ttcgttaccg tgattggcat ggaagtgatt 60

gctgcactgg cacacaaata catcatgcac ggctgggggt ggggatggca tctttcacat 120

catgaaccgc gtaaagggtgc gtttgaagtt aacgatcttt atgccgtggg ttttgctgca 180

ttatcgatcc tgctgattta tctgggcagt acaggaatgt ggccgctcca gtggattggc 240

gcaggatatga cggcgtatgg attactctat tttatgggtgc acgacgggct ggtgcatcaa 300

cgttggccat tccgctatat tccacgcaag ggctacctca aacggttgta tatggcgcac 360

cgtatgcata acgccgtcag gggcaaagaa ggttgtgttt cttttggctt cctctatgcg 420

ccgcccctgt caaaacttca ggcgacgctc cgggaaagac atggcgctag agcgggcgct 480

gccagagatg cgcagggcgg ggaggatgag cccgcatccg ggaagtaa 528

<210> 72

<211> 762

<212> DNA

<213> *Nostoc sp. PCC73102*

<400> 72

atgatccagt tagaacaacc actcagtcac caagcaaac tgactccagt actgagaagt 60

aaatctcagt ttaaggggct tttcattgct attgtcattg ttagcgcacg ggtcattagc 120

ctgagtttat tactttccct tgacatctca aagctaaaat tttggatggt attgcctggt 180

atactatggc aaacatTTTT atatacggga ttatttatta catctcatga tgccatgcat 240
ggcgtagtat ttccccaaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 300
ctttatgggc ttttaccata tcaaaaacta ttgaaaaaac attggttaca ccaccacaat 360
ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg 420
tattttcatt ttatgaaagg ttactggagt tgggggcaaa taattgcgtt gactattatt 480
tataactttg ctaaatacat actccatata ccaagtata atctaetta cttttgggtg 540
ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt 600
gaaccaatag ggggttatgt tcagcctcat tgtgccc aaa caattagccg tcctatttgg 660
tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atatcctcat 720
atttcttggg ggcagttacc agaaatttac aaagcaaaat ga 762

<210> 73

<211> 617

<212> DNA

<213> *Haematococcus pluvialis*

<400> 73

taggggtgcgg aaccaggcac gctggtttca cacctcatgc ctgtgataag gtgtggctag 60
agcgatgcgt gtgagacggg tatgtcacgg tcgactgggc tgatggccaa tggcatcggc 120
catgtctggt catcacgggc tggttgcctg ggtgaagggt atgcacatca tcatgtgcgg 180
ttggaggggc tggcacagtg tgggctgaac tggagcagtt gtccaggctg gcgttgaatc 240
agtgaggggt tgtgattggc ggttgtgaag caatgactcc gcccatattc tatttgtggg 300
agctgagatg atggcatgct tgggatgtgc atggatcatg gtagtgcagc aaactatatt 360
cacctagggc tgttggtagg atcaggtgag gccttgcaca ttgcatgatg tactcgtcat 420

gggtgtgttg tgagaggatg gatgtggatg gatgtgtatt ctgagacgta gaccttgact 480
ggaggcttga tcgagagagt gggccgtatt ctttgagagg ggaggctcgt gccagaaatg 540
gtgagtggat gactgtgacg ctgtacattg caggcagggtg agatgcactg tctcgattgt 600
aaaatacatt cagatgc 617

<210> 74

<211> 1208

<212> DNA

<213> *Haematococcus pluvialis*

<400> 74

attgtgactg atagcgagac tctgggtcga tgttatctgc ctcaacaatg gcttagaaaa 60
gaagaaacag aacaaatata gcaaggcaac gcccgtagcc taggtgatca aagactgttg 120
ggcttgtctc tgaagcttgt aggaaaggca gacgctatca tggtagagagc taagaagggc 180
attgacaagt tgccggcaaa ctgtcaaggc ggtgtacgag ctgcttgcca agtatatgct 240
gcaattggat ctgtactcaa gcagcagaag acaacatatc ctacaagagc tcacttaaaa 300
ggaagcgaac gtgccaagat tgctctgttg agtgtatata acctctatca atctgaagac 360
aagcctgtgg ctctccgtca agctagaaaag attaagagtt tttttgttga ttagtgaatt 420
tttgttttat ttatgtctga tagttcaata aagagacaac acatacaata taaaatcatt 480
gtctttaaat gttaatttag tagagtgtaa agcctgcatt tttttgtac gcataaacia 540
tgaattcacc ccgcttctgg tttttaaata attatgtcaa actagggaaa attctttttt 600
ttctcttcgt tctttttttg gcttgttgtg gagtcacagg cttgtcttca gattgataga 660
ggttgatac actcaacaga gcaatcttgg cacgttcgct tcctttttaga tgagctcttg 720
taggatatgt tgtcttctgc tgcttgagta cagatccaat tgcagcatat acttggcaag 780

cagctcgtac accgccttga cagtttgccg gcaacttgtc aatgcccttc ttagctctca 840
ccatgatagc gtctgccttt cctacaagct tcagagacaa gcccaacagt ctttgatcac 900
ctaggctacg ggcgttgcct tgctgtatct gttctgtttc ttcttttcta agccattgtt 960
gaggcagata acatcgaccc aacatcctcg agccatacta cagcataaaa ggatacgttt 1020
tctttaacag aaatttaccc ttttgttatc agcacatata aaaaaaaga aatttaagat 1080
gagtaggact tccattctct caaaaatttt attcaatcca taaatgaatt atttttggac 1140
aaaaaagaaa gattatgcct gattttctct attttttttt tttttacaac tccaccaata 1200
ctttctag 1208

<210> 75

<211> 6316

<212> DNA

<213> *Blakeslea trispora*

<220>

<221> misc_feature

<222> (2694) .. (2694)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (4263) .. (4263)

<223> n is a, c, g, or t

<400> 75

aaggatgaag aatccaactc taataaaaaat cttatggata tctttgatcg actcaaaaag 60
gctttcaatg ctattgctat taataaaaaa gagagagaga gaactatgag caaaaggact 120
ctatgccaaag atggcaaaaa ggcaccagaa acccttagtt tattattgca taatccagtc 180
gagctagtac ttctgtagct caagcttaac cgaggatctt ggaatcaact cgtctcgtca 240

ctcttgccga tgatcctaga aatggatatct atggatgtta tactaacatt gttatctttc 300

aaggcctcga agatgttatt gttgcggtga taaataggct gctatgtact gaagttgctc 360

tgtaaaatga atctagttca ctgcctactc agcaaattgg tgtttctaata gtcttttaaag 420

aaagaaaaaa agatacatat agactaccct tcctttcaag actgtaatcg agaatcggcc 480

gatggtttat tacaattaga cgctgggaat aagcaaaagg attcatcttt gtaaataaga 540

gactggtgca tatgaaagca aggatcgtat caaggaatag ttttgatcga gcatcaccag 600

caaatgctgc taatgttggc ttcttctttg cttcctgaga ttgaatggga tgtgcctaga 660

gcattgctat ttttaagtgt atactttaga tttgtgtctt tagatttgtg tcattttatt 720

tagtcaagaa agatccccct ttctctatgt atgctaagaa gaaggagcaa gaagtgtatt 780

tacaagttgg aatgagattg aaatattgta cataataata ataaaaagaa aggtagatca 840

aaaaaaatgt tctgcctatt gtaagaaatc gggaccaaca ggtgcttgat aaccagaagt 900

agcttccaat tcaggtagag gctctaggga caaatacaca attatgacag gaattttctt 960

gttgacttga acactacaag agaaacgggt cagcacaaaa tccgaaaaaa aaaagaaacg 1020

gaccattcat gtcttaccta tctagctctt tgtcttcaat tgcateccat tgctcaacca 1080

cagatacgct tcccaattga gtatattgat gaagtgttcc ctgcattttt cgcttgacta 1140

attccactac agtcacagtc ttattaatgt tttgtccttt accagtcagg ataatatgat 1200

ctttttgctt cttctatcaa aaaaataatt cttgttttga ataaaaaaaa caaatattta 1260

aagaaactac tttgatgacg gtacctggaa taactcgaga cacacatcta catatgcgtt 1320

gattttattg tggctaattc gaacctcatt ttctgctggt gggggctggt gactttcagt 1380

tgctgagacg tccttcttgc ttcttttata gtcttcact atgattttta tcaagaaagt 1440

aagtcagtga tgattgttac aagctatata tcttgaaaaa gaacagagag gtattattat 1500

cagatgcaac atgggttttct gtatcatttt catttcagtt tctctgttca aaaaaaaaaa 1560
gaacactttc tctttccact cctcaaattt tttctgctaa actcctcgca aaacatgtat 1620
ttgctttaaa ctacaagttg caattgtctg atttagcaat ttcaatatgc cttttgtgaa 1680
tccacccaaa aataaacaag tgcttgagta tacttggggt cagttcaaaa gaaagcaagc 1740
tttttttttt ctttcttggg aaagaaaaaa aaatattggt gagccatcct ttaccagcag 1800
tatgcgagct acgacatagc tgggtctaaca atgactgcaa gcaatagatc gagcttagtc 1860
tttctattgc ttcyttgttt gatctatgtt cggccttacg ctgacctatc caatactoga 1920
gataggcaac aagatttcga acagtaatga aataaatttc ggataacagt tgtggatgag 1980
gaagagaaag cgacttgaac tcgagaaaact ttgttgaaat gaaatccgac cttttacgtg 2040
atcatcatgt attatcctct ttttcttttt tttcgtagtg aattacttac tgattgcgct 2100
caagtcgcgt ctttataaag aagaaaaaaa aatattagaa ctttcaaaaa atataactga 2160
aaataaaagt gtggctcgga gagcaaatac cacatccttt gtcttgcgtt tggtaacacg 2220
gttaataagc cactataggt gaataatgat catttctgag aataaagcgc ggcttgaagc 2280
ttatatccat atcaggattc atattaggca caactcacia ttgaggttcc agaagtgcc 2340
attttttttt cctgatagcc tgtccaatta agatcaaaaa ccaactgagtt ttctctatat 2400
attttttttt ttcataattc ttaactcttc ttcctctctc tctctctctc tctctttttg 2460
gcttgcaaaa aaaatcttta gtaataccaa agaaagcaaa ccttttcctt ttcttatttc 2520
cttgcttggt ttttaatttt tgatttctct atgctttaaa taccatttc tttctttctt 2580
ctgctattac ctatcttttc attcctctcc cccctctctc tcttgggtcta taaacatcat 2640
gaagtctct tttaaaagt cgcttgacat ttatgctggt tatatacagc atcntgtgtt 2700
ttccaagtggt ttcattcttg cttttgttct ttcgattttc ctcaacactt atctactgaa 2760

cgcttcgaag caacagccca aagtgataat caaaaagggtt attgagcggg tagaagtacc 2820

aagtagagaa caacctaaat cagtcataaa gccctcctcc aagaaacact cttctcatca 2880

tcagtctgat gtcattcgcc ctcttgatga agtattgggt ttgctcgga caccgagggc 2940

cttgactgat gaagagatca tctctattgt tcaagctgggt aaaatggccc cctatgctct 3000

tgaaaagggtc ttgggcgatt tagagcgcgc tgtccatata cgtcgtgctt tgatctcccg 3060

tgactctcgt acgaaaactt tggaagacag tatgcttccc gtgaaaaact atcattatga 3120

taaagtcatg ggtgcttggt gtgaaaatgt cattgggttat atgcctattc cagtaggtgt 3180

cgcaggtaag aagttcaaca agtcgcgata ttgacaagt tgetcatcat ttctogaaaca 3240

ggtcctttgg tgattgatgg tgattctatt catattccca tggcaactac ggaagggtgt 3300

ttagttgctt ctactgccag aggttgtaaa gcaatcaatg ctggtggtgg tgccaacaca 3360

attgttggtg ctgatggtat gactcgaggt ccttggtgctg aatttcctac aatcactcgc 3420

gctgctgact gtaaacgatg gattgaacaa gaggggtgaag ctatcgtgac cgaggcattc 3480

aattcaactt ctcgttttgc tcgtgttcgt aaattgaaag ttgctcttgc cggctcgtcta 3540

gtctacatcc gtttctctac cactacaggt gatgcaatgg gcatgaacat gatctccaag 3600

ggttggtgaaa aggctttaag caagattgct gagagatata ctgatatgca gatcatttct 3660

ctttctggta actattgtac tgacaagaaa cctgctgcta tcaactggat tgaaggacgt 3720

ggtaaactctg ttgttgctga sgctgtcatc cctgggtacgg ttgtcgaaaa ggtattgaag 3780

acctctgtta gtgctttgggt tgagctgaac atctctaaaa acctggttgg ttctgctatg 3840

gctggctccg tcggtggctt taacgctcat gctgctaata ttctaactgc catttacctt 3900

gctactgggtc aagatcctgc tcaaaatgta sagagttcta actgtattac ttgatgaaa 3960

gctgtcaatg gcgaaagaga ccttcatata tcttgtacaa tgccctgtat tgaagtaggc 4020

accattgggtg gtggtactat tttgcctcct caacaagcca tgttggattt cattgggtgtg 4080

cgtgggtcctc accctaccga acctgggtgcc aatgcccgcw gccttgctcg tgttatctgt 4140

gcctctgtga tggctgggtga attgtcttta tgtgcagctt tggctgctgg tcatcttgta 4200

aaggcacaca tggctcataa tcgtaatacc actgctgctg ccgctgttgt tcctgcccct 4260

aanggcatag ttgatgtctc tacacctcct gctacacctg cagaaaagaa tgatcctatt 4320

cctggaagtt gtatcaagtc atagaattaa tattatatat atatcatata caaaaaaaaag 4380

aaaaaaaaa cactacatct atttatatatt ctccatgtac acacacacac acacatataa 4440

aaactcttta ttttccaata ttttgctttt ataaataatc ttatttcatt ctaaataaac 4500

tgtttttttt tattaatcat caaaccctgc tgagagctgt gcaatatcat ctatgttttc 4560

atgggttaac tctgggtatcg gwcgagcctc ctctgtactt gaagtttgta ggcagttttt 4620

atthaaggct gctgggtcgat catgatcatc akcaaacctg acagcatgaa gttttgactg 4680

atgagcaatt tctaataaggg cagaatctga actctttcgc ttctactat tgaccatatt 4740

gtcttttaggt ggaatgagtg aatagcgtct tgtcatatgt aacacagaat caacaatatc 4800

ctgggtgatga aactcggcca aacatagcgc ctttctcccc caacaattat aataatcaaa 4860

atgagaatga catgtacggt tttcctcgat gacaatatcc aacgtcttgt cataatcctc 4920

tgtgcgyata ccattcatct tttggaagaa cgcacggtag ctctcacaag ctgtcctcag 4980

agagttccgt gccatgtttc ccaatgctcc tggcaagtcg aaatgaagtt gtcgaatctg 5040

gcgatgtatg tctacaatgt cgctgtttc tttcattaga tcaagcattc gtgtagccca 5100

aatgatgtct atgttatgat tttctttcat tccagtaata actatagttt ctcggcaaatt 5160

cgaatgastg atggagtaaa ttcatacaaaa gtgcaagtaa tacatacagt gcttgaagaa 5220

atcttgtgta gcacgcctat attatgtaat ataggatcga ttctcgaaac tcgacataac 5280

caccaggcctt tagcaagcgt tttatttcat tcatgacaag ctattggttaa ttcytgctta 5340
ataaaacaaa atgaaaaaaaa cataccccc tcmaaactta cttcccactc ttgattggaa 5400
aaacaggatat agacgtgacg catatgtata taatcaaac actcatcagg atagggtaaa 5460
ccattgagca catcgcatcg ggtgaagaaa gtattaggag gcttgatggc tgtaggatata 5520
atagggtgcaa tatcaatacc gtaaaactca gcatttgagg attctgtagc catctccaga 5580
atccaagtac ctgtgccaca agcaacatca agcacttttag gtaagggtat acattgttgt 5640
tcttggttgtt gttgttgaca atcacttgag tctgagtttc gttttgattg ttttaatgac 5700
aataattctt ttacagggtc tgagaaatta ccgtcaaata gatacttgta aataaaatgc 5760
taaaaataaa aacaatagaa aaaaaaattg acgctcattt cattactatg gaaataactg 5820
caaaatctta ccacttgtag aagtctatct tgctcaatct catcgtttgg cagaatgtat 5880
ttattgttgt agtattgata tcttctacca ttcattgatata aactgtcgct tctaattgctc 5940
tgagggtgaag tacttgtagg tgaagggtgga agtgacgcaa ttttgctcaag cttaacagga 6000
tcctctcggc tacatgtttt ctgcatatca ggaaaatctt gtttatttga aacatcaaca 6060
gtagatgtgg tgtgatcttt ttgaaaata tcgatgcctt cctttgaaag ccttttgaaa 6120
ggctctttta acttttttga gtgagagcta cccatgatag cttatgaaga attaaaaaga 6180
aaaaagcaaa aaaaattaaa aaaaaaaaaa gtagcaaaaa attctgtcgt aattatacaa 6240
gccaatcaaa atcgaaattc atgcaaggca tagatgttca cgtggatttg atggttgatc 6300
cttttttttt gcaaga 6316

<210> 76

<211> 1170

<212> DNA

<213> *Thermus thermophilus*

<400> 76

atgaagcgcc tttccctgag ggaggcctgg ccctacctga aagacctcca gcaagatccc	60
ctcgccgtec tgctggcgtg gggccgggccc ccccccggc ttttccttcc cctgccccgc	120
ttccccctgg ccctgatctt tgaccccgag ggggtggagg gggcgctcct cgccgagggg	180
accaccaagg ccaccttcca gtaccgggccc ctctcccgcc tcacggggag gggcctcctc	240
accgactggg gggaaagctg gaaggaggcg cgcaaggccc tcaaagacct cttcctgccc	300
aagaacgtcc gcggtaccg ggaggccatg gaggaggagg cccgggcctt cttcggggag	360
tggcgggggg aggagcggga cctggaccac gagatgctcg ccctctccct gcgcctcctc	420
gggcggggccc tcttcgggaa gcccctctcc ccaagcctcg cgagcacgc ccttaaggcc	480
ctggaccgga tcatggccca gaccaggagc cccctggccc tcctggacct ggccgcccga	540
gcccgttcc ggaaggaccg gggggccctc taccgcgagg cggaagccct catcgccac	600
ccgcccctct cccaccttcc ccgagagcgc gccctgagcg aggcctgac cctcctggtg	660
gcgggcccacg agacggtggc gagcgccctc acctggtcct ttctcctcct ctcaccgcg	720
ccggactggc agaagcgggt ggccgagagc gaggaggcgg ccctcgccgc cttccaggag	780
gccctgaggc tctaccccc cgccctggatc ctcaccgga ggctggaaag gcccctcctc	840
ctgggagagg accggctccc cccgggcacc acctgggtcc tctcccccta cgtgacctag	900
aggctccact tccccgatgg ggaggcctc cgccccgagc gcttcttga ggaaagggg	960
acccttcgg ggcgctactt cccctttggc ctggggcaga ggctctgctt ggggcgggac	1020
ttcgccctcc tcgagggcc catcgctcctc agggccttct tccgccgctt ccgcctagac	1080
cccctcccct tccccgggt cctcgcccag gtcacctga ggccgaagg cgggcttccc	1140
gcgcggccta gggaggaggt gcgggcgtga	1170

<210> 77

<211> 2981

<212> DNA

<213> *Blakeslea trispora*

<400> 77

tctagaattc attccattcg aaaggatcaa cataaccaat ttaatgacta ctagctaattg	60
gatacaaata tacgcacaaa aaaagaaaga attctatgat caaagagAAC acagacacag	120
agtgatacat ttaaattggtt aagttcttat gatgttaaaa tggtaacttt attattgaat	180
taaatgcgaa tatcgttgct gctttgtact tggaaaacgt taggtaaaag ttggttaatg	240
aaagaagcag gagttgtagt atcatctctt gggaagaaat agaaaaagag gaaagtaaca	300
aagtaacaag caagacaata atagatccaa tggctttcgg tcttacgagt ttgttcagga	360
gcatacttct tttggctatc ttgtaacttt cttggtaagg gattctggcc aaagctttta	420
cagacttggt cggaagtaag cttacttcca gcaagaacga taggaacacc agtacctgga	480
tgtgtactac aaagaaaaga gaaatgagta cgtgcgttat taaaaaaaag aaaaaaagag	540
ggcaaaagta ttacctagct ccgacaaaga aaagattatc ataacggttt gtggaatcct	600
tgggtactagg tctgaaccag agaacttgga acacatcatg agaaagacca agaatagaac	660
ctctccaaag gttaaacttg ctttgccaaa cactaggatc attcacttct tcatgttcaa	720
tcaaattagc aaagttgttt actcccaaac gacgttcgat aacttcaga accatcttgc	780
gtgcacgggtt taccaactca ggataathtt ottcagcact gtttctgtc ttactcttca	840
tatggccaat tggaaccaac acaataatgg agtccttggt gggagggtgcg gcagattcat	900
caattcgaga tggaacgttg acatagaatg aagcttcaga gggcaaaccg aagtcgttga	960
aaatctcatc aaaactttcc ttgtaggctt cagccaagaa gatatttgtt acgtctaatt	1020

gaggcacctt tgttgacatg gaccaataaa acgaaataga tgatgaagtg agtttctttg 1080

aggctaattgt cttctttgtc caattgcaag gaggtaacag atggtgataa gcataaacia 1140

gatccgcatt acatacgact gcacggcctt caatgacttc tccgctttcc aaagtgcacac 1200

cggttacacg cttgtcttta tcgacagtgt taatttttagc aacaggcgat tgatatctga 1260

attcagcacc gtactttttg gaggcgatag actcaagctt ctgaacaacc atgttgaaac 1320

caccacgagg ataccagata ccttcagcaa actcgggtga ttgtaacaaa ctgtaaactg 1380

ctggagcatc ataaggcgac atactatatt ccaaaaatag aaaatagaac aatgaatatc 1440

aaaattcctt tcacttgccc tttttcacat ttctcttttc ccacccccga cgggtctcac 1500

tcattttttt ttcacccac accacgcgtt gtatgtgtac ttaccccata tacattgttt 1560

gaaaagtaaa agccatacgc attttcttgg tttggaaata tttactggct cggtcataga 1620

tcttacaaa caagtgaag cgaaagattt caggcacata ctgaagacga atcaaattccc 1680

aatgggtttc aaagttgcgc ttgatagcaa taaatgtacc ttgttcataa tggacatgtg 1740

tttccttcat gaaatccaag aatctaccaa atccaagggg accctcaata cgggtccaatt 1800

cgcccttcat cttgggttaa tcggaagaga gttgtacggc atcacgcgtc tcaaaatgaa 1860

ccttatagtt attgtcacag cgaagcaa at ccaatgatc accaatacgt tcatccaaat 1920

cagcaaatgc atcttcaaaa agcttaggca tcaaatagag tgagggaccc tgatcaaagc 1980

gatgaccatc gtgatgaatg aatgaacaac ggccaccgga aaagtcgttc ttttcaacia 2040

cagtaactcg aaaaccttca cgagcaagac gagcagcagt agcagttccg ccaataccgg 2100

caccaatgac aacaatatgc ttcttttgat cagacatgag attaaaatag ataaggaaaa 2160

gaaagtgaia agaaattcgg aagcatggca cattcttctt tttataaata catgcctgac 2220

tttctttttc catcgatatg atatatgcat atgatagata tacaagcaat cttcttcaag 2280

gagtttgaaa ttttgtcctc caggagcaaa aaaaagtttt tttttatata tgtttgtaca 2340
caagaatagt taccaatttg ctttggctct acgtgctgca agtttatatc gttttcaatt 2400
tctttgtctt tacattttct ttgtccttta tctttcctca tttagtcttt gggagaatta 2460
ggaaaaggga gcggaaaggt aagaaatgct tgcgtatttt actaatcgg caaacatcca 2520
atttggcaaa cagcagcctg tgcaacgctc tcgagatgac agtatctttg attacactct 2580
aatctcgcgt gacccgacca aaaagagcga acaaagaaat aatcttgtgc attcgaatat 2640
gatggaagat tttttccccc ttattctaaa tgttgacata gcgtgtatgt tatataaaca 2700
aaaagaaatt gtacaaactt tcttttcttc tctttttatt ttatctctat gtcaatactc 2760
acttatctgg aatttcatct ctactatata ctacctgtcc ttgcggcatt gtgttggtg 2820
ctaaagccgt ttcactcaca gcaagacaat ctcaagtata aatttttaat gttgatggcc 2880
gcctctaccg catcgatttg ggacaattat atcgtttatc atcgcgcttg gtggtactgt 2940
cctacttgtg ttgtggctgt cattggctat gtacctctag a 2981

<210> 78

<211> 1749

<212> DNA

<213> Blakeslea trispora

<400> 78

atgtctgatc aaaagaagca tattgttgtc attggtgccg gtattggcgg aactgctact 60
gctgctcgtc ttgctcgtga aggttttcga gttactgttg ttgaaaagaa cgacttttcc 120
ggtggccggt gttcattcat tcatcacgat ggtcatcgct ttgatcaggg tccctcactc 180
tatttgatgc ctaagctttt tgaagatgca tttgctgatt tggatgaacg tattggtgat 240
catttggatt tgcttcgctg tgacaataac tataaggttc attttgacga cggatgatgcc 300

gtacaactct cttccgattt aaccaagatg aagggcgaat tggaccgtat tgaggggtccc 360
cttggatttg gtagattctt ggatttcatg aaggaaacac atgtccatta tgaacaaggt 420
acatttattg ctatcaagcg caactttgaa accatttggg atttgattcg tcttcagtat 480
gtgcctgaaa tctttcgctt gcacttgttt ggtaagatct atgaccgagc cagtaaatat 540
ttccaaacca agaaaatgcg tatggctttt acttttcaaa caatgtatat gggatatgtcg 600
ccttatgatg ctccagcagt ttacagtttg ttacaataca ccgagtttgc tgaaggatc 660
tggtatcctc gtggtggtt caacatggtt gttcagaagc ttgagtctat cgctccaaa 720
aagtacggtg ctgaattcag atatcaatcg cctgttgcta aaattaacac tgtcgataaa 780
gacaagcgtg taaccggtgt cactttggaa agcggagaag tcattgaagc cgatgcagtc 840
gatatgtaatg cggatcttgt ttatgcttat caccatctgt tacctccttg caattggaca 900
aagaagacat tagcctcaaa gaaactcact tcatcatcta tttcgtttta ttggtccatg 960
tcaacaaagg tgcctcaatt agacgtacac aatatcttct tggctgaagc ctacaaggaa 1020
agttttgatg agattttcaa cgacttcggt ttgccctctg aagcttcatt ctatgtcaac 1080
gttccatctc gaattgatga atctgccga cctcccaaca aggactccat tattgtgttg 1140
gttccaattg gccatatgaa gagtaagaca ggaaacagtg ctgaagaaaa ttatcctgag 1200
ttggtaaacc gtgcacgcaa gatggttctg gaagttatcg aacgtcgttt gggagtaaac 1260
aactttgcta atttgattga acatgaagaa gtgaatgatc ctagtgtttg gcaaagcaag 1320
tttaaccttt ggagaggttc tattcttggt ctttctcatg atgtgttcca agttctctgg 1380
ttcagacctt gtaccaagga ttccacaaac cgttatgata atcttttctt tgtcggagct 1440
agtacacatc caggtactgg tgttcctatc gttcttgctg gaagtaagct tacttccgac 1500
caagtctgta aaagctttgg ccagaatccc ttaccaagaa agttacaaga tagccaaaag 1560

aagtatgctc ctgaacaaac tcgtaagacc gaaagccatt ggatctatta ttgtcttgct 1620
tgttactttg ttactttcct ctttttctat ttcttcccaa gagatgatac tacaactcct 1680
gcttctttca ttaaccaact ttacctaac gttttccaag taaaagcag caacgatatt 1740
cgcatttaa 1749

<210> 79

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 79

ccgatggcga cgacggaagg ttgtt 25

<210> 80

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 80

catgttcacg cccattgcat cacct 25

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/000100

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N1/15 C12N15/80 C12P23/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98/46772 A (BOVENBERG ROELOF ARY LANS ;GIST BROCADES BV (NL); SELTEN GERARDUS) 22 October 1998 (1998-10-22) the whole document	1-7,9, 14, 19-23, 31-33,35
Y	VAN HEESWIJCK R ET AL: "HIGH FREQUENCY TRANSFORMATION OF MUCOR WITH RECOMBINANT PLASMID DNA" CARLSBERG RESEARCH COMMUNICATIONS, vol. 49, no. 7, 1984, pages 691-702, XP009031431 ISSN: 0105-1938 the whole document	1-3,7, 14, 19-21, 31-34
	--- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *8* document member of the same patent family

Date of the actual completion of the international search

27 May 2004

Date of mailing of the international search report

29/06/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Vogt, T

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP2004/000100

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>NAVARRO EUSEBIO ET AL: "Overexpression of the crgA gene abolishes light requirement for carotenoid biosynthesis in Mucor circinelloides"</p> <p>EUROPEAN JOURNAL OF BIOCHEMISTRY, vol. 267, no. 3, February 2000 (2000-02), pages 800-807, XP002282392</p> <p>ISSN: 0014-2956</p> <p>the whole document</p>	<p>1-3,7, 14, 19-21, 31-34</p>
Y	<p>NAVARRO E ET AL: "A negative regulator of light-inducible carotenogenesis in Mucor circinelloides"</p> <p>MGG MOLECULAR GENETICS AND GENOMICS, vol. 266, no. 3, November 2001 (2001-11), pages 463-470, XP002282393</p> <p>ISSN: 1617-4615</p> <p>the whole document</p>	<p>1-3,7, 14, 19-21, 31-34</p>
Y	<p>SKORY C D: "Homologous recombination and double-strand break repair in the transformation of Rhizopus oryzae."</p> <p>MGG MOLECULAR GENETICS AND GENOMICS, vol. 268, no. 3, November 2002 (2002-11), pages 397-406, XP002282394</p> <p>ISSN: 1617-4615 (ISSN print)</p> <p>the whole document</p>	<p>1-3,7, 14, 19-21, 31-34</p>
A	<p>US 6 413 736 B1 (JOLLY SETSUKO OMATA ET AL) 2 July 2002 (2002-07-02)</p> <p>the whole document</p>	
A	<p>US 2002/051998 A1 (SCHMIDT-DANNERT CLAUDIA ET AL) 2 May 2002 (2002-05-02)</p> <p>the whole document</p>	
A	<p>LEE P C ET AL: "Metabolic engineering towards biotechnological production of carotenoids in microorganisms."</p> <p>APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 60, no. 1-2, October 2002 (2002-10), pages 1-11, XP002282395</p> <p>ISSN: 0175-7598</p>	
A	<p>ZHIQIANG A ET AL: "New cosmid vectors for library construction, chromosome walking and restriction mapping in filamentous fungi"</p> <p>GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, ELSEVIER SCIENCE PUBLISHERS, BARKING, GB, vol. 176, no. 1-2, 17 October 1996 (1996-10-17), pages 93-96, XP004070214</p> <p>ISSN: 0378-1119</p>	

-/--

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP2004/000100

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,Y	<p>NICOLAS FRANCISCO E ET AL: "Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes."</p> <p>EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, vol. 22, no. 15, 1 August 2003 (2003-08-01), pages 3983-3991, XP002282396</p> <p>ISSN: 0261-4189 (ISSN print)</p> <p>the whole document</p>	1-3,31

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP2004/000100

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9846772	A	22-10-1998	AU 7642298 A	11-11-1998
			BR 9808859 A	01-08-2000
			CN 1257546 T	21-06-2000
			WO 9846772 A2	22-10-1998
			EP 0979294 A2	16-02-2000
			JP 2001518798 T	16-10-2001
			PL 336345 A1	19-06-2000
			US 6432672 B1	13-08-2002
<hr/>				
US 6413736	B1	02-07-2002	US 6015684 A	18-01-2000
			US 5922560 A	13-07-1999
			US 5466599 A	14-11-1995
			US 2003049241 A1	13-03-2003
			AU 688280 B2	12-03-1998
			AU 5008693 A	08-11-1994
			EP 0708604 A1	01-05-1996
			JP 8508885 T	24-09-1996
			NO 954052 A	11-12-1995
			NZ 255742 A	26-05-1997
			WO 9423594 A1	27-10-1994
<hr/>				
US 2002051998	A1	02-05-2002	AU 2257401 A	18-06-2001
			EP 1238068 A1	11-09-2002
			WO 0142455 A1	14-06-2001
<hr/>				

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2004/000100

A. KLASIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
 IPK 7 C12N1/15 C12N15/80 C12P23/00

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 7 C12N C12P

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, BIOSIS

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	WO 98/46772 A (BOVENBERG ROELOF ARY LANS ;GIST BROCADES BV (NL); SELTEN GERARDUS) 22. Oktober 1998 (1998-10-22) das ganze Dokument	1-7,9, 14, 19-23, 31-33,35
Y	VAN HEESWIJCK R ET AL: "HIGH FREQUENCY TRANSFORMATION OF MUCOR WITH RECOMBINANT PLASMID DNA" CARLSBERG RESEARCH COMMUNICATIONS, Bd. 49, Nr. 7, 1984, Seiten 691-702, XP009031431 ISSN: 0105-1938 das ganze Dokument	1-3,7, 14, 19-21, 31-34



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderscher Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderscher Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

Z Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

27. Mai 2004

Absendedatum des internationalen Recherchenberichts

29/06/2004

Name und Postanschrift der internationalen Recherchenbehörde
 Europäisches Patentamt, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Vogt, T

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2004/000100

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	<p> NAVARRO EUSEBIO ET AL: "Overexpression of the crgA gene abolishes light requirement for carotenoid biosynthesis in Mucor circinelloides" EUROPEAN JOURNAL OF BIOCHEMISTRY, Bd. 267, Nr. 3, Februar 2000 (2000-02), Seiten 800-807, XP002282392 ISSN: 0014-2956 das ganze Dokument </p>	<p> 1-3,7, 14, 19-21, 31-34 </p>
Y	<p> NAVARRO E ET AL: "A negative regulator of light-inducible carotenogenesis in Mucor circinelloides" MGG MOLECULAR GENETICS AND GENOMICS, Bd. 266, Nr. 3, November 2001 (2001-11), Seiten 463-470, XP002282393 ISSN: 1617-4615 das ganze Dokument </p>	<p> 1-3,7, 14, 19-21, 31-34 </p>
Y	<p> SKORY C D: "Homologous recombination and double-strand break repair in the transformation of Rhizopus oryzae." MGG MOLECULAR GENETICS AND GENOMICS, Bd. 268, Nr. 3, November 2002 (2002-11), Seiten 397-406, XP002282394 ISSN: 1617-4615 (ISSN print) das ganze Dokument </p>	<p> 1-3,7, 14, 19-21, 31-34 </p>
A	<p> US 6 413 736 B1 (JOLLY SETSUKO OMATA ET AL) 2. Juli 2002 (2002-07-02) das ganze Dokument </p>	
A	<p> US 2002/051998 A1 (SCHMIDT-DANNERT CLAUDIA ET AL) 2. Mai 2002 (2002-05-02) das ganze Dokument </p>	
A	<p> LEE P C ET AL: "Metabolic engineering towards biotechnological production of carotenoids in microorganisms." APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, Bd. 60, Nr. 1-2, Oktober 2002 (2002-10), Seiten 1-11, XP002282395 ISSN: 0175-7598 </p>	
A	<p> ZHIQIANG A ET AL: "New cosmid vectors for library construction, chromosome walking and restriction mapping in filamentous fungi" GENE: AN INTERNATIONAL JOURNAL ON GENES AND GENOMES, ELSEVIER SCIENCE PUBLISHERS, BARKING, GB, Bd. 176, Nr. 1-2, 17. Oktober 1996 (1996-10-17), Seiten 93-96, XP004070214 ISSN: 0378-1119 </p>	

-/-

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
PCT/EP2004/000100

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
P,Y	<p>NICOLAS FRANCISCO E ET AL: "Two classes of small antisense RNAs in fungal RNA silencing triggered by non-integrative transgenes."</p> <p>EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, Bd. 22, Nr. 15, 1. August 2003 (2003-08-01), Seiten 3983-3991, XP002282396 ISSN: 0261-4189 (ISSN print) das ganze Dokument -----</p>	1-3,31

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2004/000100

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9846772 A	22-10-1998	AU 7642298 A	11-11-1998
		BR 9808859 A	01-08-2000
		CN 1257546 T	21-06-2000
		WO 9846772 A2	22-10-1998
		EP 0979294 A2	16-02-2000
		JP 2001518798 T	16-10-2001
		PL 336345 A1	19-06-2000
		US 6432672 B1	13-08-2002
US 6413736 B1	02-07-2002	US 6015684 A	18-01-2000
		US 5922560 A	13-07-1999
		US 5466599 A	14-11-1995
		US 2003049241 A1	13-03-2003
		AU 688280 B2	12-03-1998
		AU 5008693 A	08-11-1994
		EP 0708604 A1	01-05-1996
		JP 8508885 T	24-09-1996
		NO 954052 A	11-12-1995
		NZ 255742 A	26-05-1997
		WO 9423594 A1	27-10-1994
US 2002051998 A1	02-05-2002	AU 2257401 A	18-06-2001
		EP 1238068 A1	11-09-2002
		WO 0142455 A1	14-06-2001